Enabling Collaborative Geo-visual Analytics

Gustavo A. Garcia-Chapeton,

Frank O. Ostermann, Rolf A. de By & Menno-Jan Kraak

ESRI USER CONFERENCE

Outline

- Motivation
- Methodology
- Collaborative systems and techniques
- Research challenges
- Conclusions

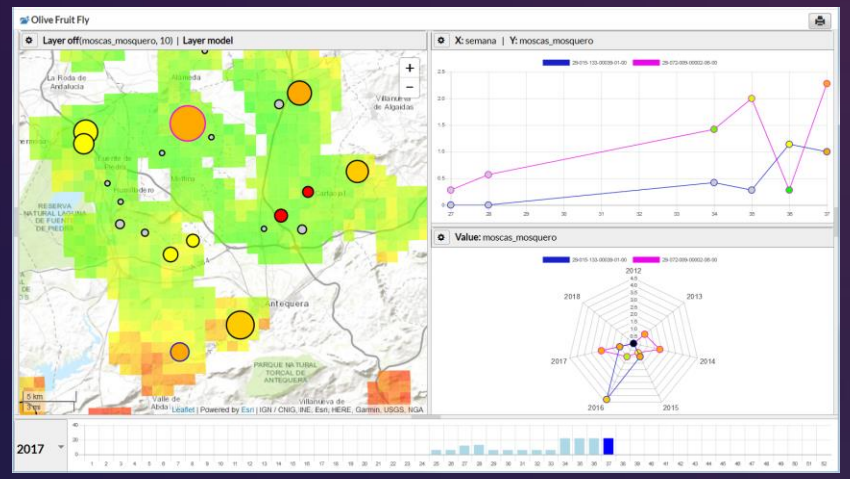
Motivation

Geo-DBs

The abundance of geo-data presents a unique opportunity to better understand natural and man-made processes.

Motivation

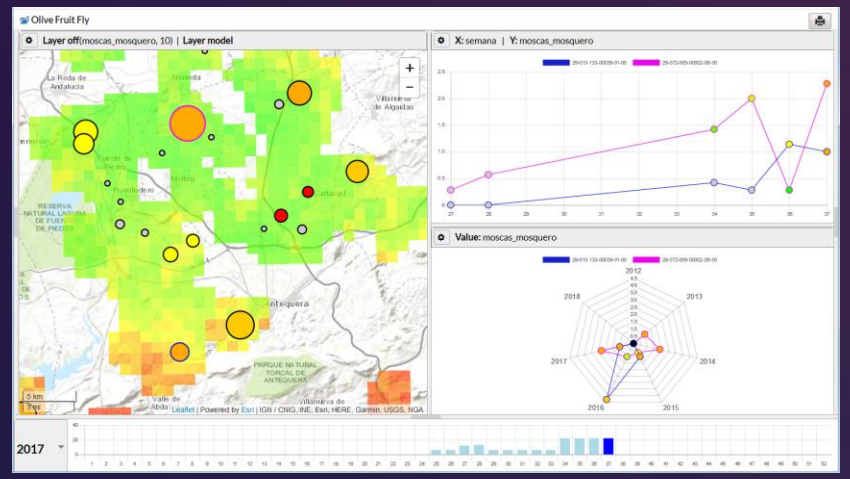
Geo-DBs



Geo-visual analytics (GVA) enables a synergy between human analytical skills and computer storage and processing power to make sense of large and complex geo-data sets.

Motivation

Geo-DBs



Analytical problems are complex, ill-defined and broad in scope, and require to be addressed by multi-disciplinary teams. However, most GVA environments are single-user environments that offer limited support for collaborative work, as a consequence, collaboration remains a challenge for research in GVA.

Methodology

Systematic review following the guidelines proposed by Kitchenham and Charters (2007).

Planning

- Define objectives
- Define information sources
- Define information to be extracted

Conducting

- Obtain information sources
- Extract, organize and synthesize information

Reporting

- Prepare a document with the results of the review
  - *Article - Enabling Collaborative Geo-Visual Analytics: systems, techniques and research challenges*

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering (Technical Report No. EBSE-2007-01). Keele, UK: Keele University.

Methodology — research objectives

1. Identify and describe GVA systems that support collaborative work
2. Identify and describe collaborative techniques in these systems
3. Identify research challenges to effectively support collaborative work in GVA and propose strategies to address these

Methodology – information sources

Information sources

Databases
- ACM Digital Library, GeoBase, IEEEXplore, Science direct, Scopus, Springer Link, and Web of Science

Search
- Keywords: (collaborative OR cooperative) AND (“geovisual analytics” OR “geospatial visual analytics” OR geoanalytics)

Inclusion
- Published between 2004 and 2017
- Journals, proceedings, and book chapters
- English

Exclusion
- Duplicated papers
- Irrelevant papers

Methodology — extracted information

1. Collaborative GVA (CGVA) systems
   1. Supported scenarios
   2. Supported techniques
   3. Deployment options
   4. Supported devices

2. Collaborative techniques
   1. Characteristics
   2. Advantages and limitations
   3. Complementary techniques

3. Claims of research challenges

Summary of results

Identified data sources (papers)

- Total 124
- Unique 99
- Relevant 28

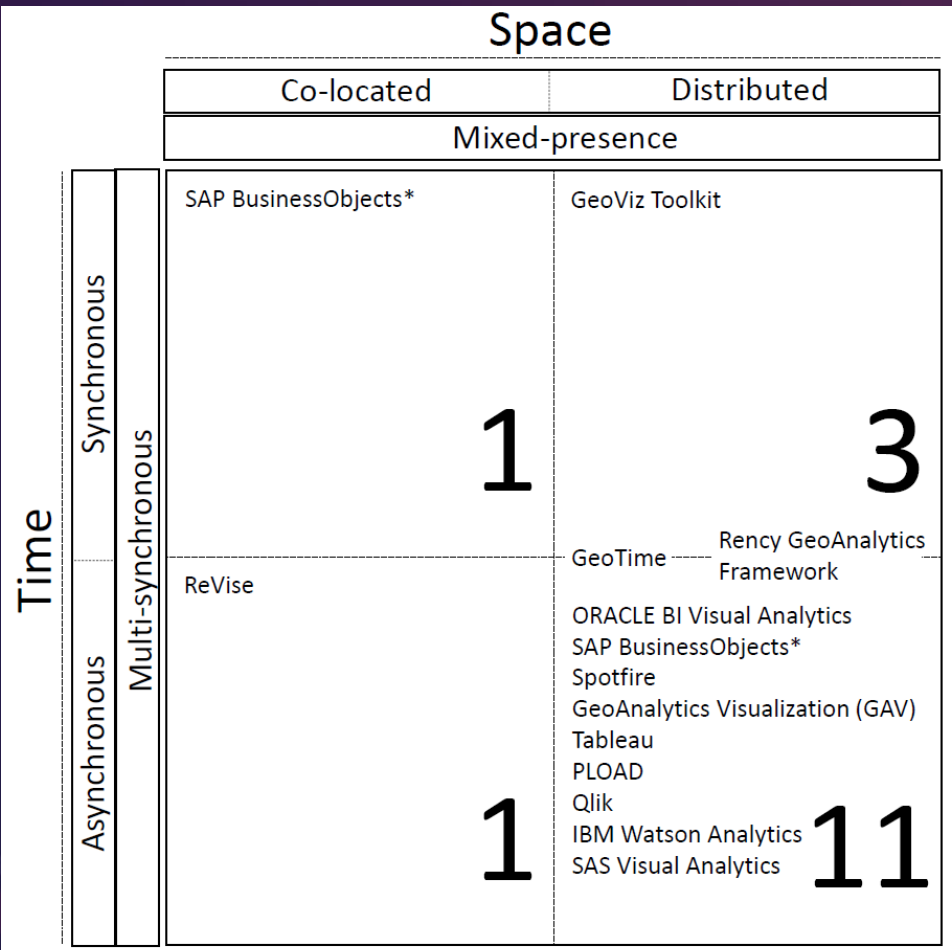
Information extracted

- Description of 13 CGVA systems
- Description of 6 collaborative techniques
- Description of 3 research challenges

Note: To identify systems not mentioned in the literature, the search was extended to the web using Google Search, which resulted in the identification of 3 systems.

Systems — collaborative scenarios

- Most common collaborative scenario: asynchronous distributed
  - 11 systems
  - Promotes participation by removing the constraints of space and time
- Only 3 out of 13 systems support hybrid collaborative scenarios



Collaborative scenario**	No	%
Synchronous co-located	1	8%
Synchronous distributed	3	23%
Asynchronous co-located	1	8%
Asynchronous distributed	11	85%
Hybrid collaborative scenario	**No**	**%**
Mixed-presence	0	0%
Multi-synchronous	2	15%
Synchronous co-located and asynchronous distributed	1	8%

No = Number of systems that support the scenario
% = Percentage of systems that support the scenario (out of 13 systems)

* SAP BusinessObjects appears twice, it was done to avoid confusion by placing it in a diagonal between synchronous co-located and asynchronous distributed.

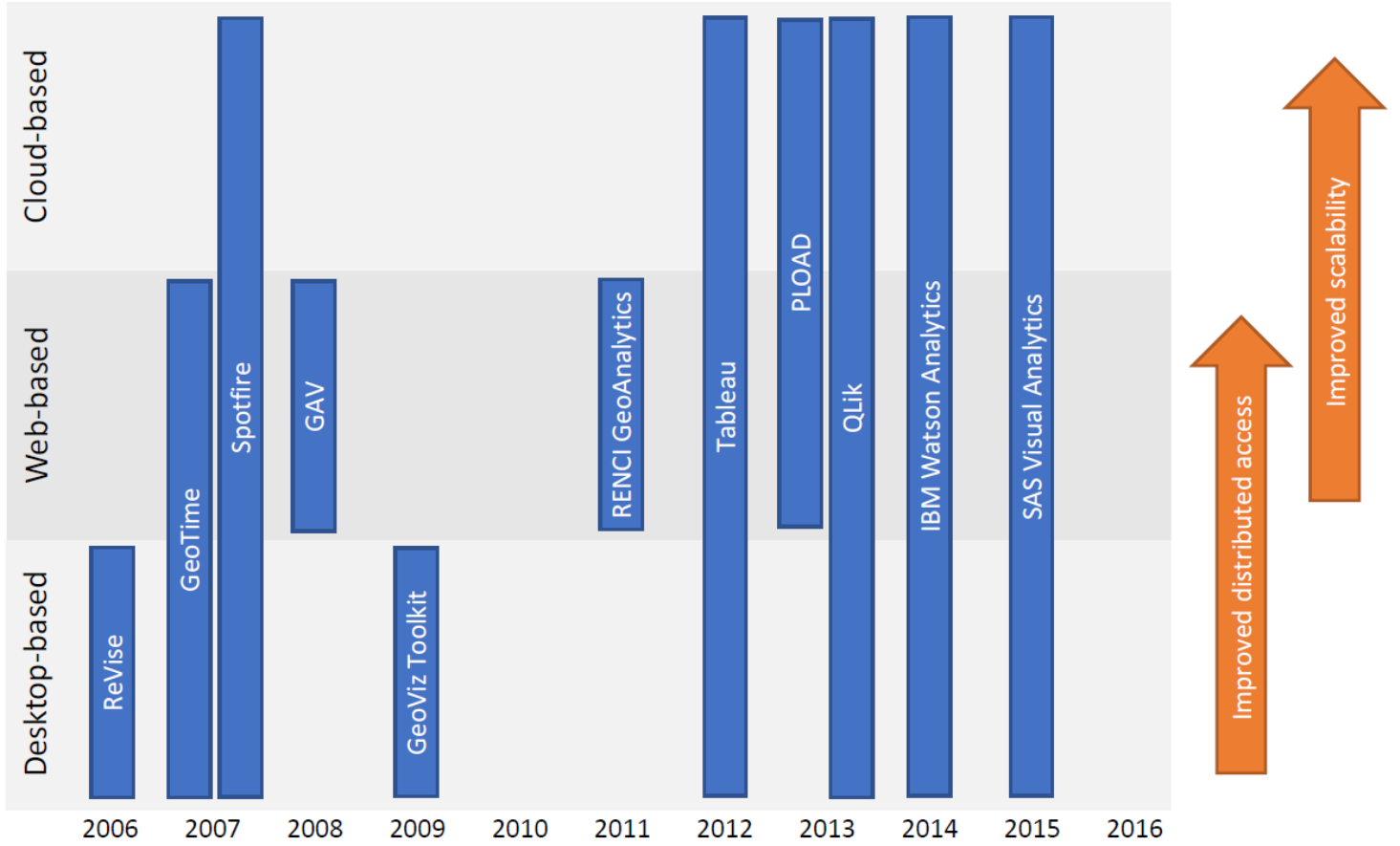
** Collaborative scenarios are not mutually exclusive, for this reason the column **No** do not sum up to 13 and **%** to 100%.

Systems — deployment options

GVA environments are increasingly using cloud-based deployment

1. Distributed access to the system enabled by the Internet, and
2. Flexible and scalable storage capacity and processing power, enabling users to work from thin clients

Web interfaces enable cross-device support in 62% of the systems.



Note: The location in this timeline is based in the first reference to a collaborative feature in the systems.

Systems and techniques

Most popular techniques:

1. Snapshot (11)
2. Storytelling (8)
3. Annotation (7)

Storytelling + snapshot + annotation:

1. Describes the analysis through a story
2. Documents relevant observations using snapshot and annotation
3. Snapshot captures the context, and annotation-specific aspects
4. Analytic results are ready for communication

	Annotation	Discussion board	Instant messaging	Interaction history	Snapshot	Storytelling	Techniques / system
ORACLE BI Visual Analytics					✓	✓	2
SAP Business Objects		✓			✓	✓	3
ReVise	✓			✓	✓		3
GeoTime	✓				✓	✓	3
Spotfire	✓	✓			✓		3
GeoAnalytics Visualization Framework					✓	✓	2
GeoViz Toolkit			✓		✓		2
RENCI GeoAnalytics Framework	✓		✓				2
Tableau	✓				✓	✓	3
PLOAD	✓						1
QLIK					✓	✓	2
IBM Watson Analytics		✓			✓	✓	3
SAS Visual Analytics	✓				✓	✓	3
System / technique	7	3	2	1	11	8	13

Techniques — co-occurrence and occurrence in collaborative scenarios

Highest co-occurring techniques:

1. Snapshot and storytelling (8)
2. Snapshot and annotation (5)

Occurrence of techniques in collaborative scenarios

1. No discernible patterns
2. Only snapshot occurs in all the scenarios

	a							b			
	Annotation	Discussion board	Instant messaging	Interaction history	Snapshot	Storytelling		Synchronous co-located	Synchronous distributed	Asynchronous co-located	Asynchronous distributed
Annotation	-	-	-	-	-	-		0	2	1	6
Discussion Board	1	-	-	-	-	-		1	0	0	3
Instant Messaging	1	0	-	-	-	-		0	2	0	1
Interaction history	1	0	0	-	-	-		0	0	1	0
Snapshot	**5**	3	1	1	-	-		1	2	1	9
Storytelling	3	2	0	0	**8**	-		1	1	0	8

Techniques — some similarities and differences

Discussion board and instant messaging

1. Based on exchange of text messages
2. Instant messaging is directed and private, limiting its usefulness
3. Discussion board is public and offers better organization (threads)

Snapshot and interaction history

1. Both document the evolution of analysis process
2. Snapshot: documents and restores discrete states on-demand
3. Interaction history: documents interactions and state changes automatically, and allows to review the analysis as a continuous process

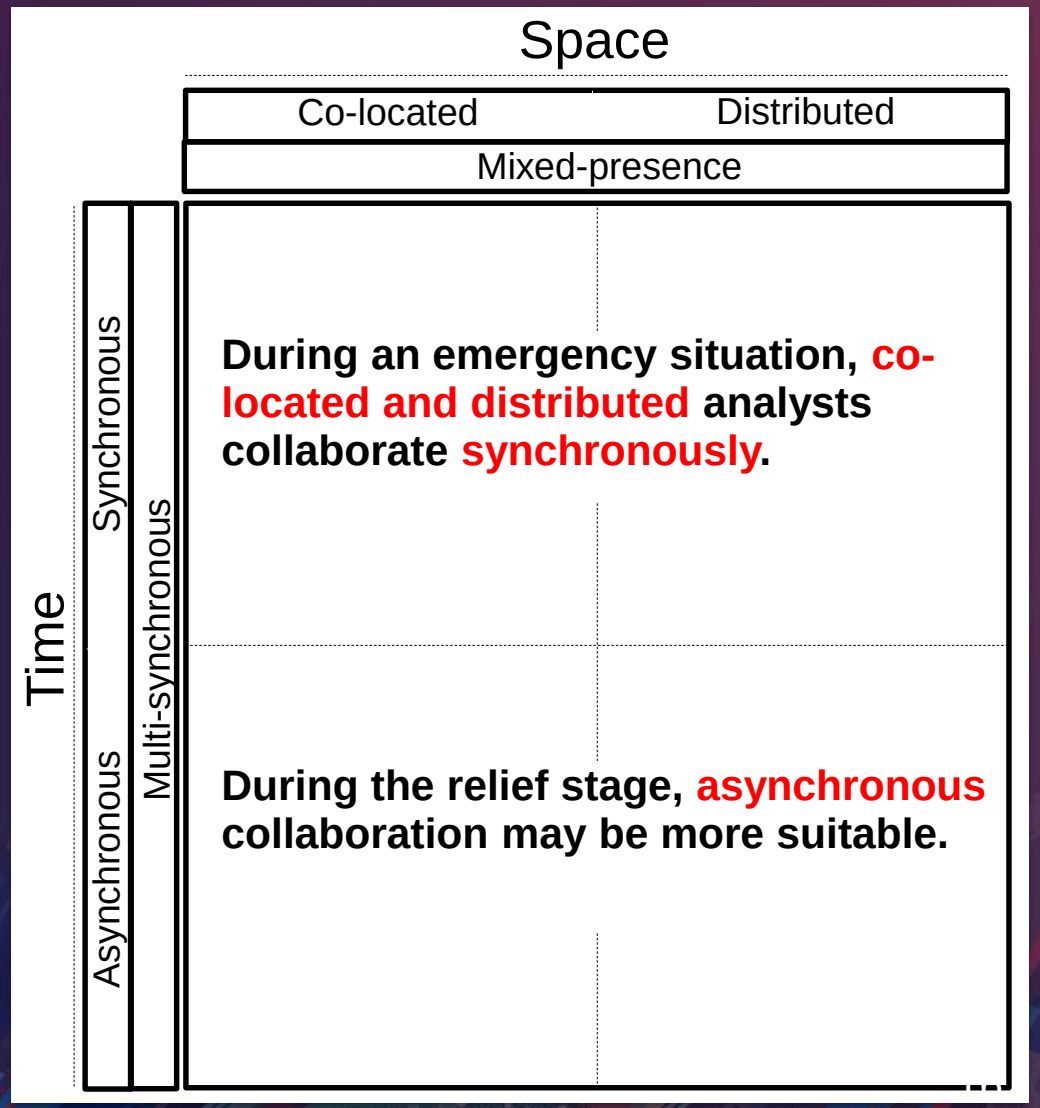
Challenge 1: hybrid collaborative scenarios

An analysis effort comprises different tasks, each of which may benefit or even require several collaborative scenarios

To address this challenge:

1. Assess advantages and disadvantages of the scenarios and their suitability for specific types of task
2. Design mechanisms that properly integrate contributions regardless the collaborative scenario
3. Special attention is required for scenarios in which analysts may work offline

Example



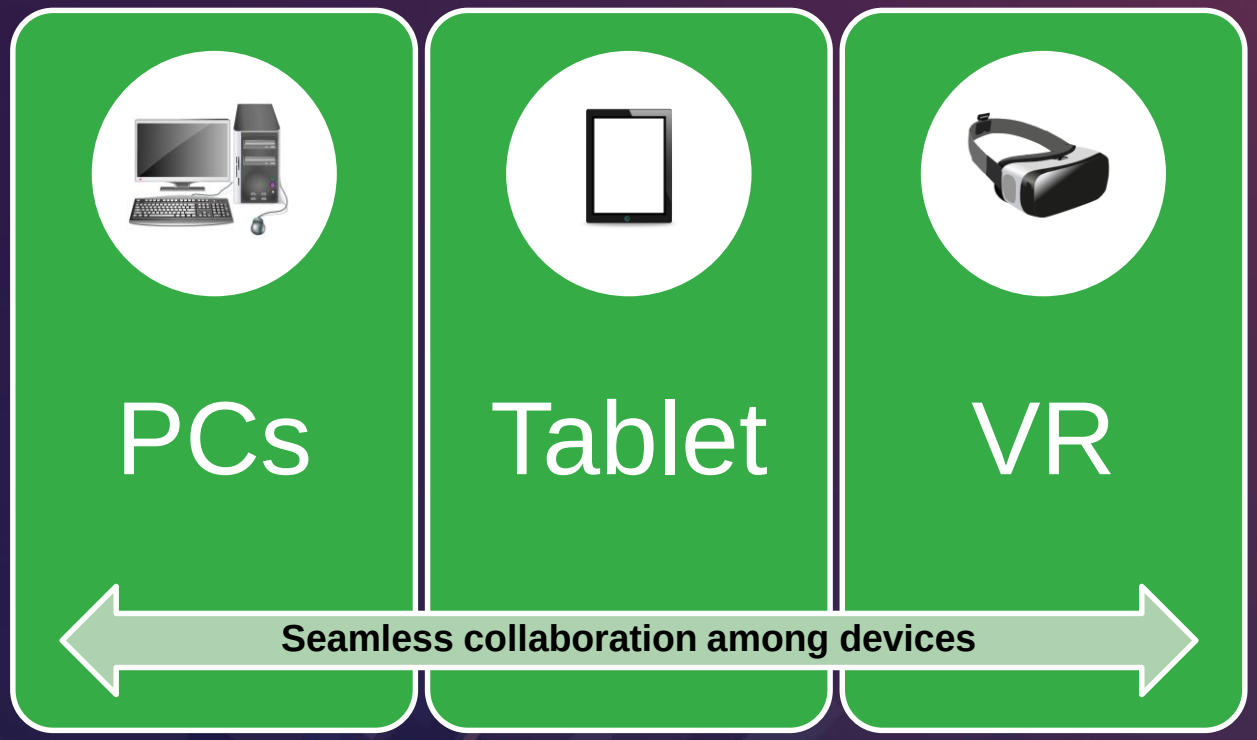
Challenge 2: cross-device collaboration

Potential of cross-device collaboration:

1. Provide a more flexible analysis workflow
2. Enable actors from diverse backgrounds to participate without requiring specialized hardware

Effective support for cross-device collaboration means:

1. Take advantage of the unique characteristics of each type of device
2. Enables seamless transition between devices
3. Ensure responsiveness regardless the device in use



Challenge 3: time-critical and long-term analysis

Time-critical

Lasts for short period, minimization of undesirable consequences requires timely delivery of results.

- Examples: natural disasters, terrorist attacks and cyber-attacks

Special attention

- Conflicting interactions
- Awareness of analysis status
- Timely communication of results

Long-term

Extends for a much longer time span and aims to generate understanding and/or enable strategic decisions.

- Examples: climate change, urban dynamics and species conservation

Special attention

- Summary of progress
- Parallel analysis projects and multiple working hypotheses
- Communication of partial and final results

Conclusions

1. CGVA is aiming to reach a broader audience
   - Asynchronous distributed collaboration: eliminates constraint of space and time
   - Cloud-based deployment: improves accessibility and scalability
   - Multiple devices: eliminates the need for specialized hardware

2. Snapshot, storytelling and annotation provides a flexible and effective working environment
   - Iterative knowledge generation process
   - Enables analysts to combine independent and collaborative work
   - Results can be immediately communicated to a broader audience

3. CGVA is moving towards effective support of multi-disciplinary and cross-domain collaborative analysis, but to materialize this potential, the identified pressing challenges need to be addressed.

Enabling Collaborative Geo-visual Analytics

Gustavo A. Garcia-Chapeton (g.a.garciachapeton@utwente.nl),

Frank O. Ostermann, Rolf A. de By & Menno-Jan Kraak

ESRI USER CONFERENCE

Identified papers

TABLE 1 Electronic databases used to identify papers for the systematic review

Source	URL	Papers	Unique results	After paper screening
ACM digital library	dl.acm.org	19	11	4
GeoBase	www.engineeringvillage.com	3	3	2
IEEEXplore	ieeexplore.ieee.org	10	10	2
Science direct	www.sciencedirect.com	12	12	3
Scopus	www.scopus.com	32	22	7
Springer Link	link.springer.com	39	34	6
Web of Science	apps.webofknowledge.com	9	6	4
Total	–	124	99	28

The column "Papers" accounts for the total of papers identified in a database, the column "Unique results" accounts for non-duplicated papers, and the column "After paper screening" accounts for papers that contribute to address the review objectives.

Advantages and limitations of the identified techniques

TABLE 3 Advantages and limitations of the identified collaborative techniques

Technique	Advantages	Limitations
Annotation	• Enables analysts to point at, describe, and bound features of interest in the data products • Can carry semantics that link the annotation with the underlying data	• Lack of guidelines to regulate its use may lead to an overload of irrelevant contributions
Discussion board	• Enables topic-centered discussion among geographically distributed analysts • Topics are organized in threads	• Synthesized discussion results is not trivial
Instant messaging	• Enables discussion among geographically distributed analysts	• Private discussions may lead to lack of awareness of others' work and to fragmentation of the known information • Discussion board is more flexible and better organized
Interaction history	• Documents the analysis as a continuous process automatically • The interaction logs can be stored and accessed based on different models • Allows the analysis process to be reviewed and extended	• An interaction history may require editing before it can be disseminated • Snapshot offers an alternative to document the analysis process as discrete states and is deemed sufficient in most use cases
Snapshot	• Allows discrete states of the analysis process to be stored on demand • Stored states can be reconstructed for further analysis • Can be applied to independent visual products or the whole analytical environment	• Unlike interaction history, snapshot cannot reconstruct the interactions that led to the stored states
Storytelling	• Organized in chapters • Supports a flexible analysis process by allowing the story to be updated • Specific focus on communication of analytical results • Effective, engaging, and easy to understand for specialists and laypersons	• Doesn't incorporate identification of individual's contributions • Doesn't offer provenance of the story