APPENDIX
Commercial Visual Analytics Systems – Advances in the Big Data Analytics Field

Michael Behrisch, Dirk Streeb, Florian Stoffel, Daniel Seebacher, Brian Matejek Stefan Hagen Weber, Sebastian Mittelstädt, Hanspeter Pfister, Daniel Keim

Abstract—Five years after the first state-of-the-art report on Commercial Visual Analytics Systems we present a reevaluation of the Big Data Analytics field. We build on the success of the 2012 survey, which was influential even beyond the boundaries of the InfoVis and Visual Analytics (VA) community. While the field has matured significantly since the original survey, we find that innovation and research-driven development are increasingly sacrificed to satisfy a wide range of user groups. We evaluate new product versions on established evaluation criteria, such as available features, performance, and usability, to extend on and assure comparability with the previous survey. We also investigate previously unavailable products to paint a more complete picture of the commercial VA landscape. Furthermore, we introduce novel measures, like suitability for specific user groups and the ability to handle complex data types, and undertake a new case study to highlight innovative features. We explore the achievements in the commercial sector in addressing VA challenges and propose novel developments that should be on systems' roadmaps in the coming years.

Index Terms—System Comparison, Commercial Landscape, Visual Analytics Research, Advances, Development Roadmap.

✦

CONTENTS

1 SURVEY ANSWERS

In this appendix we report an aggregate/ summary of the conducted survey questionnaire for our state-of-the-art report. Specifically, we give the answers to all yes/no questions for the survey. Nevertheless, it should be noted that a wide free text answers helped us to understand and reason on the vendors responses. These comment fields are excluded here (partly due to the wish of some vendors.)

- *M. Behrisch, B. Matejek, H. Pfister are with Harvard University, Cambridge, USA*
  E-mail: {behrisch, bmatejek, pfister}@g.harvard.edu
- *D. Streeb, F. Stoffel, D. Seebacher, D. Keim are with University of Konstanz, Germany.*
  E-mail: {dirk.streeb, florian.stoffel, daniel.seebacher, daniel.keim}@uni-konstanz.de.
- *S. H. Weber, S. Mittelstädt are with Siemens AG, Corporate Research Germany.*
  E-mail: {stefan_hagen.weber, sebastian.mittelstaedt}@siemens.com.

2 DATA HANDLING AND -MANAGEMENT

Data Management, ETL and Preprocessing	Tableau	Advizor	JMP	Cognos	Jaspersoft	Spotfire	Visual Analytics	PowerBI
Data Sources — Import								
CSV files	✓	✓	✓	✓	✓	✓	✓	✓
Excel-files (.xls/.xlsx)	✓	✓	✓	✓	✓	✓	✓	✓
JSON APIs (REST)	✓	✓	✓		✓			✓
JSON-files	✓	✓	✓	✓	✓		✓	✓
XML (REST)	✓	✓	✓		✓			✓
XML files		✓	✓	✓	✓		✓	✓
SQL-Databases (e.g. ODBC/Postgres)	✓	✓	✓	✓		✓	✓	✓
NOSQL-Databases (e.g. MongoDB)	✓			✓	✓	✓	✓	✓
Directory of text/image/video			✓				✓	✓
OPC (UA)							✓	
SAP	✓			✓		✓		✓
Salesforce	✓	✓		✓	✓	✓		✓
Data Sources — Export								
CSV	✓	✓	✓	✓	✓	✓	✓	✓
Excel	✓		✓	✓	✓	✓	✓	✓
JSON				✓	✓		✓	
XML				✓	✓		✓	
SQL			✓			✓	✓	
OPC (UA)								
Writeback								
To data source	✓		✓			✓		
ETL — Complex relational queries/functions								
By scripting		✓	✓	✓	✓	✓	✓	✓
Via visual query interfaces	✓		✓	✓	✓	✓	✓	✓
Via wizards		✓	✓		✓	✓	✓	✓
ETL before data is loaded in-memory								
Predefined data	✓	✓	✓	✓	✓	✓	✓	✓
Data transformations implementable by user (e.g. in JAVA, R)	✓	✓	✓		✓	✓	✓	✓
Table operations like joins	✓	✓	✓	✓	✓	✓	✓	✓
Preprocessing — (Semi-) Automatic Data Preprocess.								
Imputation of values (e.g.	✓	✓	✓			✓	✓	✓
rmalisation of dimensions		✓				✓	✓	✓
Filtering by values	✓	✓	✓	✓	✓	✓	✓	✓
Sub-sampling		✓	✓				✓	✓
Outlier detection		✓	✓				✓	✓

Fig. 1. **Data Handling and -Management:** The commercial VA systems are comparable in terms of their offered input and export functionalities, their ETL- and preprocessing functionality

3 AUTOMATIC ANALYSIS

		Automatic Processing	Tableau	Advizor	JMP	Cognos	Jaspersoft	Spotfire	Visual Analytics	PowerBI
Automatic Algorithms	Data Mining	Classification	✓	✓	✓	✓		✓	✓	✓
		Clustering	✓	✓	✓	✓		✓	✓	✓
		Outlier detection	✓	✓	✓	✓		✓	✓	✓
		Association mining/ Frequent			✓	✓		✓		
		Regression analysis	✓	✓	✓	✓		✓	✓	✓
		Time-series prediction	✓		✓	✓		✓	✓	✓
		Automatic aggregations/	✓	✓				✓	✓	✓
	Statistics	Percentiles	✓	✓	✓	✓		✓	✓	✓
		Moments			✓			✓	✓	
		Normality-test (e.g. Shapiro-			✓	✓		✓	✓	
		Goodness of fit (e.g. R^2, Chi^2)		✓	✓	✓		✓	✓	✓
		Model comparison (e.g. AIC,			✓	✓		✓	✓	✓
		Influence measures (e.g. Cook's		✓	✓	✓		✓	✓	✓
		Hypothesis tests (e.g. t-test,			✓	✓		✓	✓	✓
User Support	Analysis/Macro Definition	Yes, in a scripting console		✓	✓			✓	✓	✓
		Yes, via a wizard		✓				✓		✓
		Yes, via a drag-and-drop			✓	✓				✓
		Yes, by recording and reapplying workflows			✓					✓
	Pattern Querying	Visual Query Interface (e.g.,		✓	✓			✓	✓	✓
		Regular Expressions	✓	✓	✓			✓		✓
		Free hand sketching	✓	✓						✓
	Support Long-Running Tasks	Yes, sampling-based computations	✓	✓	✓			✓		✓
		Yes, prediction methods	✓	✓						
		Yes, incremental calculations		✓	✓				✓	

Fig. 2. **Automatic Analysis:** Although many vendors reported their support for many (sophisticated) data mining algorithms we found that they (partially) rely on their implemented analysis bridges to R, Java or other scripting languages.

4 COMPLEX DATA TYPES

Supported Data Types and Streaming Support		Tableau	Advizor	JMP	Cognos	Jaspersoft	Spotfire	Visual Analytics	PowerBI
Data Types	Numeric with error	✓		✓			✓	✓	✓
	Numeric intervals		✓	✓			✓	✓	✓
	Complex time series	✓	✓	✓				✓	✓
	Item Sets	✓	✓		✓			✓	✓
	Relational/Networks	✓	✓					✓	✓
	Geographic positions	✓	✓	✓	✓	✓	✓	✓	✓
	Geographic traces	✓					✓		✓
	Geographic areas	✓	✓		✓		✓	✓	✓
	Text	✓	✓	✓	✓	✓	✓	✓	✓
	String-like (e.g. DNA)		✓						✓
	Image			✓	✓	✓	✓	✓	✓
	Video				✓				
Streaming — Dataset	Updates as approaching							✓	✓
	Updates by fixed time intervals		✓	✓	✓	✓	✓		
	Updates by fixed amounts			✓					✓
Streaming — Visualization	Manual updates	✓	✓	✓		✓	✓		✓
	Periodic updates	✓	✓	✓					✓
	Dynamic near real-time	✓							✓

Fig. 3. **Complex Data Types:** Additionally, to the data types table already described in the paper we show here also how streaming data is supported for dataset updates and the resulting visualization updates. Esp. *Tableau* and *PowerBI* are here setting new standards.

5 VISUALIZATION

	Visualization	Tableau	Advizor	JMP	Cognos	Jaspersoft	Spotfire	Visual Analytics	PowerBI
Non-Standard Visualizations	Parallel coordinate plots	✓	✓	✓	✓		✓	✓	✓
	Sunburst charts	✓			✓				✓
	Treemaps	✓	✓	✓	✓	✓	✓	✓	✓
	Node-link diagrams	✓	✓	✓	✓			✓	✓
	Matrix-based visualizations			✓	✓	✓	✓	✓	✓
	Geographic maps	✓	✓	✓	✓	✓	✓	✓	✓
	Heatmaps	✓	✓	✓	✓	✓	✓	✓	✓
	Dense Pixel-based	✓	✓		✓		✓	✓	✓
	Glyph-based small multiples	✓	✓	✓	✓				✓
	Horizon-graphs	✓	✓		✓				✓
Output Devices	Large displays/ Powerwall	✓	✓	✓		✓	✓	✓	✓
	Multi-display desktop	✓	✓	✓			✓	✓	
	Desktop	✓	✓	✓	✓	✓	✓	✓	✓
	Tablet	✓	✓	✓	✓	✓	✓	✓	✓
	Phone	✓	✓	✓	✓	✓	✓	✓	
	Simultaneous combinations of the above	✓	✓	✓		✓	✓	✓	✓
Reporting	Journals/ Linear histories		✓	✓			✓	✓	✓
	Tree-like histories			✓			✓		✓
	Data-flow graphs/ Visual analysis workflow depiction		✓				✓		✓

Fig. 4. **Visualization:** All commercial VA systems support standard visualizations, like bar-, pie and line charts. For the more sophisticated chart functionality we found out that many vendors who reported their support for more (sophisticated) visualizations (partially) rely on their implemented visualization bridges to D3 and JavaScript, or R, Java.

6 USER-GUIDANCE, PERCEPTION, COGNITION

User-Guidance, Perception and Cognition			Tableau	Advizor	JMP	Cognos	Jaspersoft	Spotfire	Visual Analytics	PowerBI
User Guidance	**General**	Visualization Suggestions	🟠		🟠			🟠	🟠	🟠
		Parameter Suggestions			🟠			🟠		🟠
	Data Preparation	Yes, by scripting		🔵	🔵	🔵	🔵	🔵	🔵	🔵
		Yes, via visual query interfaces	🔵		🔵	🔵	🔵	🔵	🔵	🔵
		Yes, via wizards		🔵	🔵	🔵	🔵	🔵	🔵	🔵
	Scripting, Reoccuring Tasks	Yes, in a scripting console		🟠	🟠			🟠	🟠	🟠
		Yes, via a wizard		🟠				🟠		🟠
		Yes, via a drag-and-drop interface			🟠	🟠				🟠
		Yes, by recording and reapplying workflows			🟠					🟠
	Pattern Retrieval	Visual Query Interface (e.g., WYSIWYG Data Transformation Pipeline)		🔵	🔵			🔵	🔵	🔵
		Regular Expressions	🔵	🔵	🔵			🔵		🔵
		Free hand sketching	🔵	🔵						🔵
Support		Getting-started online tutorials	🟠	🟠	🟠	🟠	🟠	🟠	🟠	🟠
		Advanced online tutorials	🟠	🟠	🟠	🟠	🟠	🟠	🟠	🟠
		Online forum	🟠	🟠	🟠	🟠	🟠	🟠	🟠	🟠
		Online wiki or other Q/A platform	🟠	🟠	🟠	🟠	🟠	🟠	🟠	🟠
		Interactive help function in system		🟠	🟠	🟠	🟠	🟠	🟠	🟠
		Extensive usage manuals/books	🟠		🟠		🟠	🟠	🟠	🟠
		Email support	🟠	🟠	🟠	🟠	🟠	🟠	🟠	🟠
		24/7 phone support	🟠		🟠	🟠	🟠	🟠	🟠	
		Workday 9/5 phone support	🟠	🟠		🟠	🟠	🟠	🟠	
		In-house support	🟠	🟠	🟠	🟠	🟠	🟠	🟠	
		In-house workshops	🟠	🟠	🟠	🟠	🟠	🟠	🟠	🟠
		Workshops at local retailer	🟠			🟠	🟠	🟠	🟠	🟠
		User conference	🟠	🟠	🟠	🟠	🟠	🟠	🟠	🟠
		Developer conference	🟠	🟠		🟠	🟠	🟠	🟠	🟠

Fig. 5. **User-Guidance, Perception, Cognition:** This subfield can be categorized into user guidance for pattern retrieval, scripting and data preparation. Most interestingly, more or less sophisticated Show Me buttons become standard, reflecting the importance of Quality-Metric driven exploration of datasets. Additionally, we report here on the vendors' offered support channels.

7 INFRASTRUCTURE

		Infrastructure	Tableau	Advizor	JMP	Cognos	Jaspersoft	Spotfire	Visual Analytics	PowerBI
General	Operating System	Windows	✓	✓	✓	✓	✓	✓	✓	✓
		Mac OS	✓		✓		✓			
		Linux				✓	✓	✓	✓	
		iOS	✓	✓			✓	✓	✓	✓
		Android	✓				✓		✓	✓
	Architecture	Stand alone	✓	✓	✓		✓	✓		✓
		Client/Server on premise	✓	✓		✓	✓	✓	✓	
		Cloud on premise (customer site)	✓	✓		✓	✓	✓	✓	✓
		Cloud at product company site	✓			✓		✓	✓	✓
		Cloud in internet (e.g. Amazon S3, Azure)	✓	✓		✓	✓	✓	✓	✓
Performance	Single PC	Yes, multi-CPU/core supported	✓	✓	✓	✓	✓	✓	✓	✓
		Yes, single-GPU comp. supported	✓					✓	✓	
		Yes, multi-GPU computation is supported						✓	✓	
	Cloud	Yes, computation can be deferred to dedicated compute clusters/nodes	✓			✓		✓	✓	✓
Extendability	How	SDK	✓			✓	✓	✓		✓
		Mashups	✓			✓	✓	✓		✓
		Integration of other software (e.g. R, D3)	✓		✓	✓	✓	✓		✓
		Integrated scripting environment			✓	✓		✓	✓	✓
	What	Import data sources	✓		✓	✓	✓	✓	✓	✓
		Merge data sources	✓		✓	✓	✓	✓	✓	✓
		Bi-directional interfaces	✓			✓		✓	✓	✓
		Preparation/ Cleaning of data	✓		✓	✓		✓	✓	✓
		Interactive visualisations	✓		✓	✓	✓	✓		✓
		Machine learning	✓		✓			✓	✓	✓
		Statistical analysis	✓		✓			✓	✓	✓
		Result output	✓		✓	✓	✓	✓	✓	✓

Fig. 6. **Infrastructure:** Additionally, to the already presented Table in the paper we show here how and what can be extended for our tested commercial VA systems. Esp. *PowerBI* sets here new standards by allowing users without visualization or programming experience to download add-ins for extending the visualization and analytic capabilities of their product.

8 TESTED SYSTEMS

System	Version	Tested in
QlikView	12	Feb. 2017
Spotfire	7.8	Feb. 2017
Tableau	10.2	Feb. 2017
SAS Visual Analytics	7.3	Feb. 2017
JMP Pro	13	Feb. 2017
Advizor	6.8	Feb. 2017
Lumira	1.31	Feb. 2017
Microsoft Power BI	2.56	Mar. 2018

Fig. 7. The table provides an overview of the tested systems, their tested product versions, and the date when the system was tested.