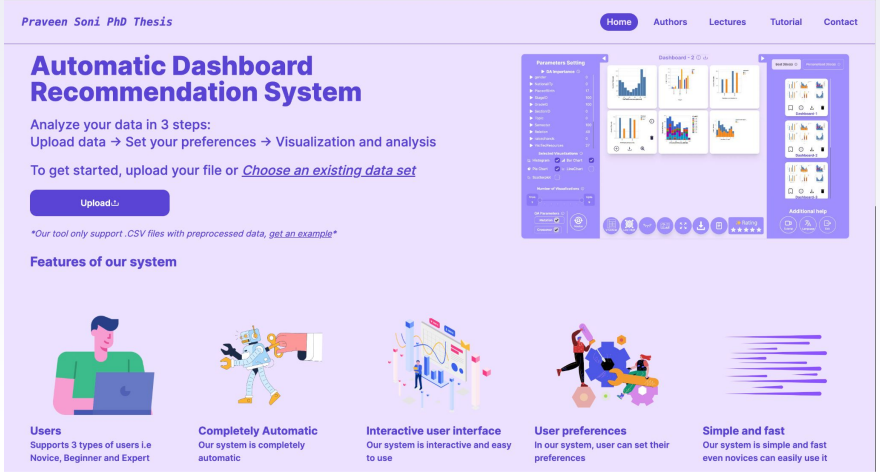
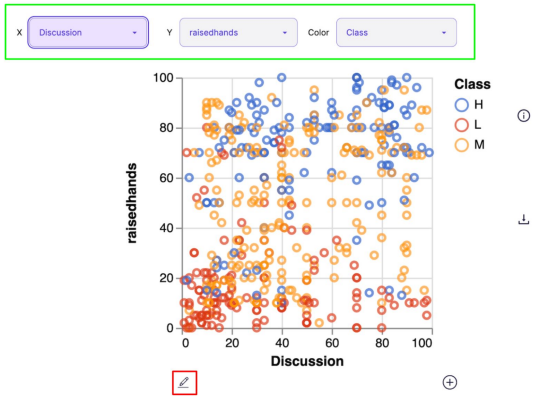
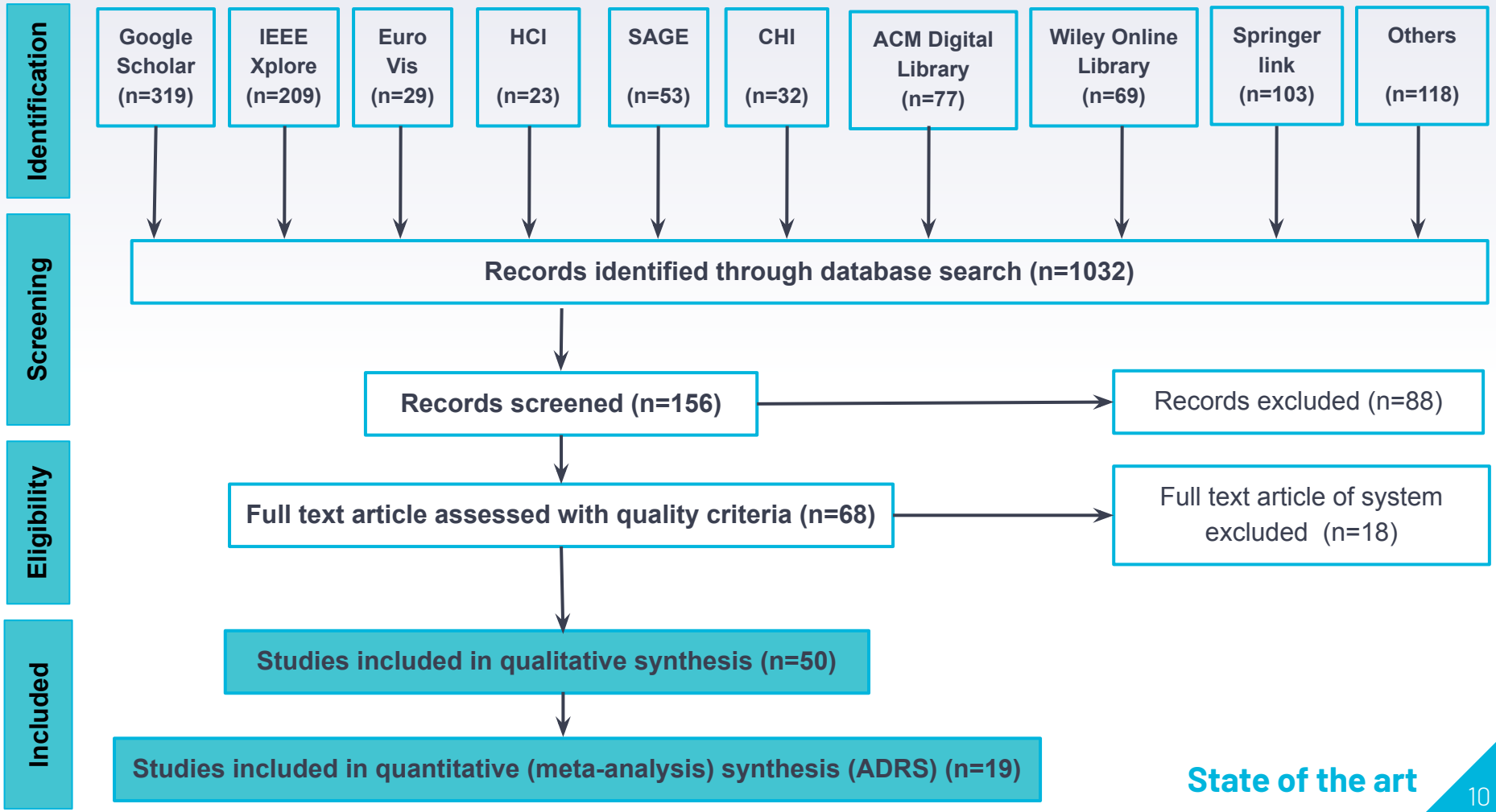
Overview

1. Introduction
2. State of the art
3. Genetic Algorithm (GA) for Dashboard generation
4. SimpleVis: User Interface
5. System Evaluation: Comparative study
6. Conclusion and Perspective





PRISMA


Identification
Screening
Eligibility
Included
Google Scholar (n=319)
IEEE Xplore (n=209)
Euro Vis (n=29)
HCI (n=23)
SAGE (n=53)
CHI (n=32)
ACM Digital Library (n=77)
Wiley Online Library (n=69)
Springer link (n=103)
Others (n=118)
Records identified through database search (n=1032)
Records screened (n=156)
Records excluded (n=88)
Full text article assessed with quality criteria (n=68)
Full text article of system excluded (n=18)
Studies included in qualitative synthesis (n=50)
Studies included in quantitative (meta-analysis) synthesis (ADRS) (n=19)

Systems Analysis

Table 3: Overview of the studied ADRSs.

System	Year and Reference	Considered vis.	Suggestion method	Utility functions	Layout	User feedback	User Interface	License
ScagExplorer	2014 (Dang and Wilkinson, 2014)	SP (1)	SD, GO	9, IN	OL, OS	None	SM	NF
DashBot	2021 (Da Col et al., 2021)	T, BC, SP, PC (4)	BB, SD, ML, SiD	3, IN	IS, IG	DF	SM	NF
CCP	2020 (Chen et al., 2020)	T, LC, BC, M, MV (14)	BB, T	None	PT, MS, ML	BF	DD, SM	NF
Exploration Views	2011 (Elias and Bezerianos, 2011)	PC, BC, T, MV (NF)	T, DM	None	PT, MS, MLa	None	DD, SM	C
DynSpace	2018 (El Meseery et al., 2018)	HT, BC, SP (3)	DM	None	IS, IG, MS, MLa	None	DD, SM	NF
VizDeck	2012 (Key et al., 2012)	HT, BC, SP, LC, PC, M (7)	SD, BB, ML	None	IS, IG	DF	SM	C
MultiVision	2021 (Wu et al., 2021)	SP, LC, BC, AC, PC (5)	BB, ML, SiD	4	IS, IG, MS, MLa	None	DD, SM	OS
Tableau (Show Me)	2007 (Mackinlay et al., 2007)	T, BC, SP, GC, MV (14)	UIAs, SI	(2) NF	IS, IG, MS, MLa	None	DD, SM	C
Power BI	2015 (Microsoft, 2015)	LC, T, BC, PC, M, SP, MV (21)	UIAs, SI	NF	IS, IG, MS, MLa	None	DD, SM	C
DeepEye	2018 (Luo et al., 2018)	BC, LC, PC, SP (4)	BB, ML	3, IN	IS, IG	None	SM	F
Draco	2018 (Moritz et al., 2018)	MT (NF)	VQL, ML	None	None	None	None	C
Zenvisage	2016 (Siddiqui et al., 2016)	HM, SP, HT, LC, MV (NF)	VQL	None	IS, IG	None	DD, SM	OS
Qlik Sense	2013 (QlikTech International, 2013)	LC, G, PC, PT, HT, MV (NF)	UIAs	NF	IS, IG, MS, MLa	None	DD, SM	C
Voyager 2	2017 (Wongsuphasawat et al., 2017)	HT, SP, LC, MT, MV (NF)	VQL	2	IS, IG, MS, MLa	None	DD, SM	C
QuickSight	2016 (Amazon Web Services, 2016)	BC, T, SP, HT, MV (NF)	UIAs	NF	IS, IG, MS, MLa	None	DD, SM	C
Sisense	2004 (Sisense Inc, 2004)	SP, PC, HT, LC, MV (NF)	UIAs	NF	IS, IG, MS, MLa	None	DD, SM	C
"Engaging Dashboards"	2019 (Aksu et al., 2019)	SP, HT, HM, FP, MV (13)	DM	None	IS, IG	None	SM	C
Datashot	2019 (Diamond and Mattia, 2017)	BC, AC, M, SD, PC (14)	BB, SD, ML	2, IN	IS, IG	None	SM	OS
Calliope	2020 (Zeng et al., 2021)	BC, LC, PC, AC, BP (12)	BB, SD, GO	1, IN	IS, IG	None	SM	OS

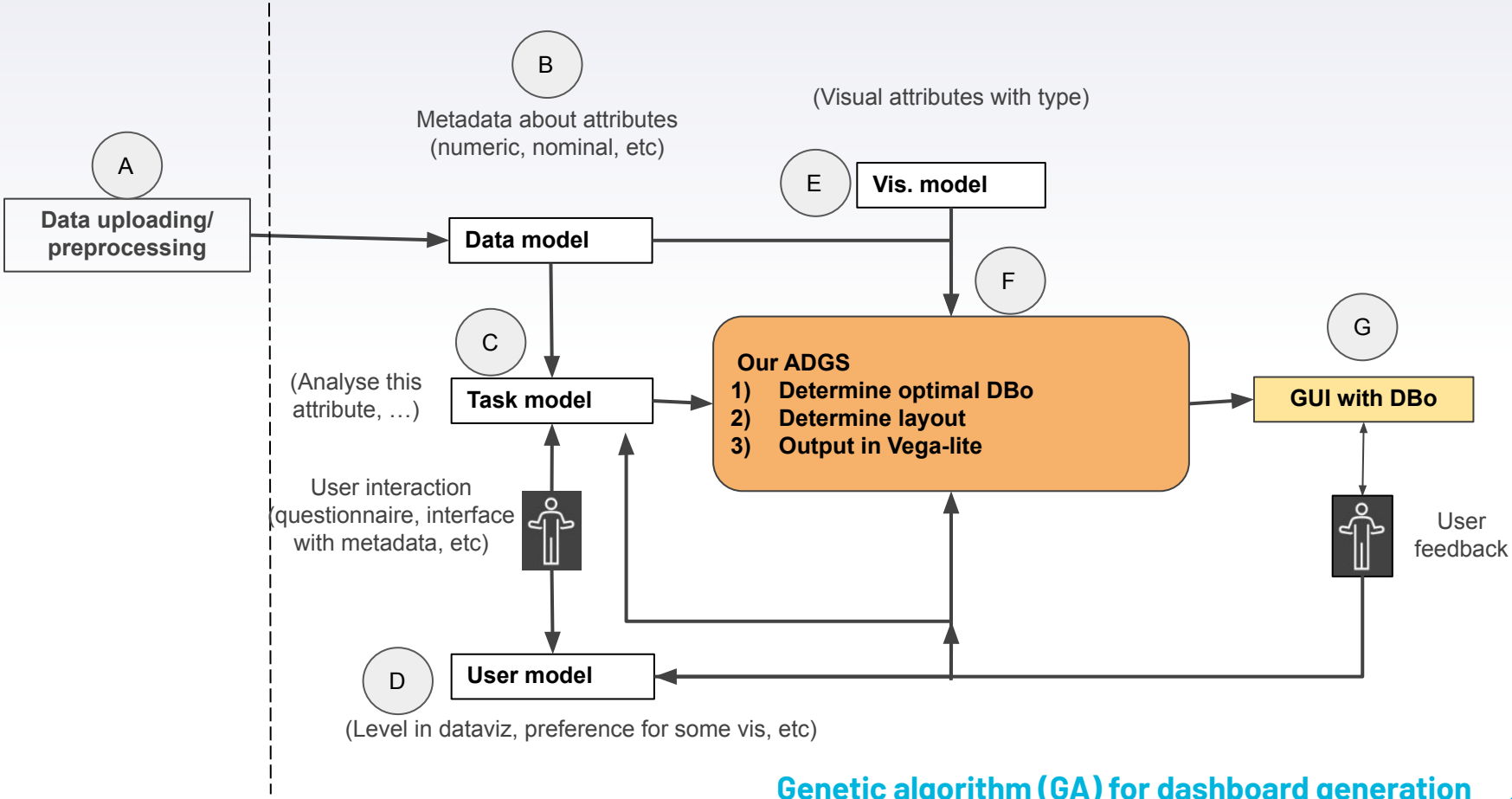
Existing Systems Evaluation

The evaluation was a demonstration of functionalities on use cases:

- **7 systems perform system testing**
- **4 systems perform comparative study**
- **1 system perform machine based evaluation (DeepEye)**
- **5 systems are commercial:** they do not provide evaluation results: except Tableau
- **3 systems do not perform any kind of evaluation:** Draco, DashBot and ScagExplorer

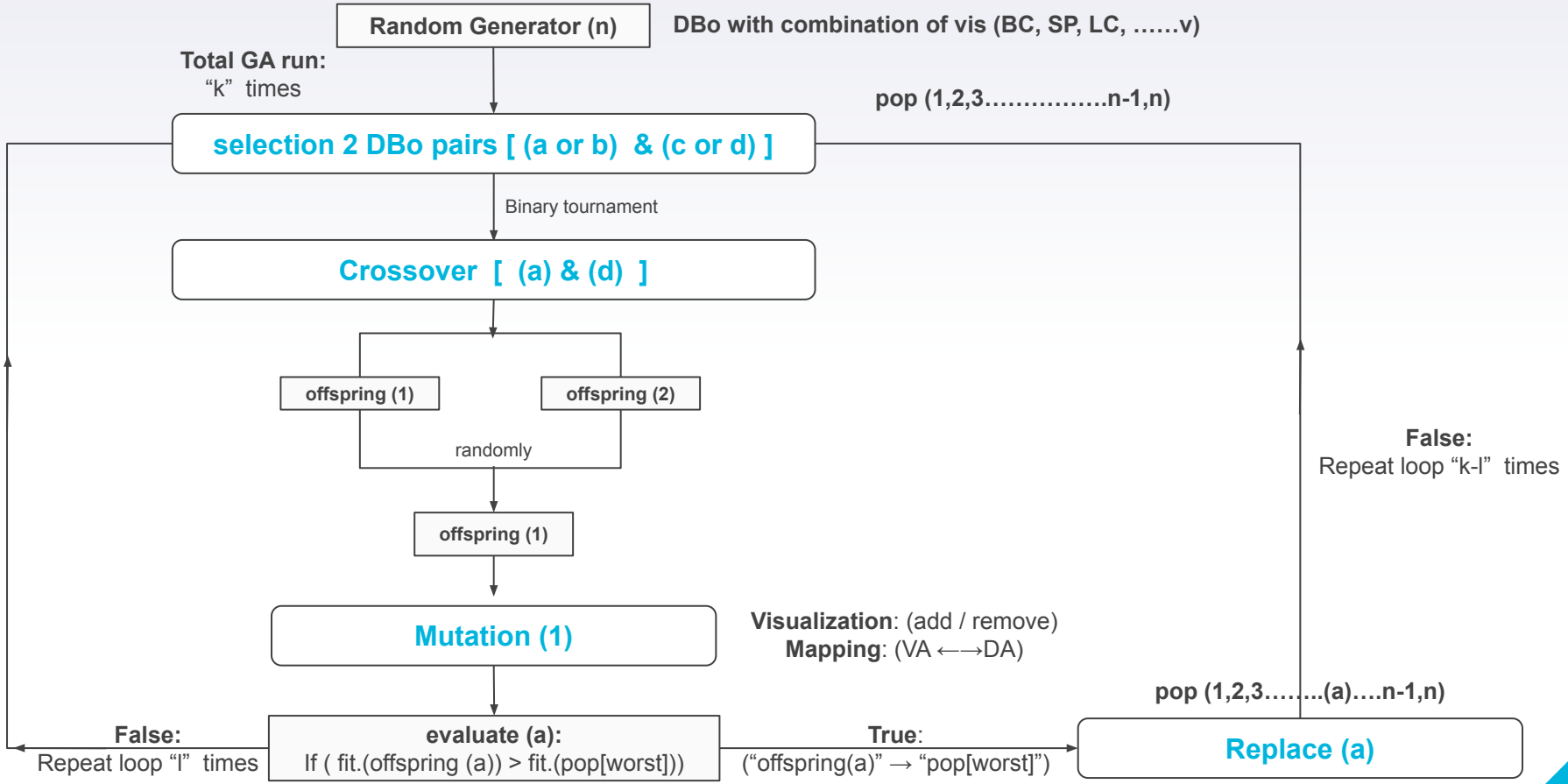
System	Type	U	Mode	Findings
CCP	Testing	20	Online	Completion time, Ease of use
Expl.View	Testing	15	Offline	Ease of use, chart creation and customization
DynSpace	Testing	8	Offline	Selecting dimensions preferred
VizDeck	Comp.	32	Offline	Completion time
MultiVision	Testing	12		Useful, convenient, easy to learn
Zenvisage	Comp.	12	Offline	Completion time, accuracy, usability and satisfaction
Voyager 2	Comp.	16	Offline	Supports focused analysis, related views + wildcards helpful
Eng. Dash.	Testing	3	Online	Dashboard improvement
Tableau	Testing	10	Online	Show Me + Add to Sheet useful
DataShot	Testing	10	Offline	Effective data fact extraction, visual design, interactions
Calliope	Comp.	16	Offline	Logicality, memorability, comprehension, engagement, dissemination

Our System Overview



Genetic algorithm (GA) for dashboard generation

Genetic Algorithm

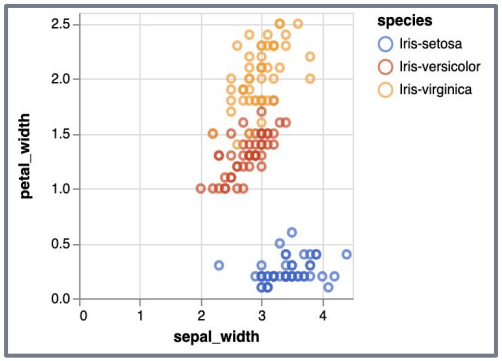


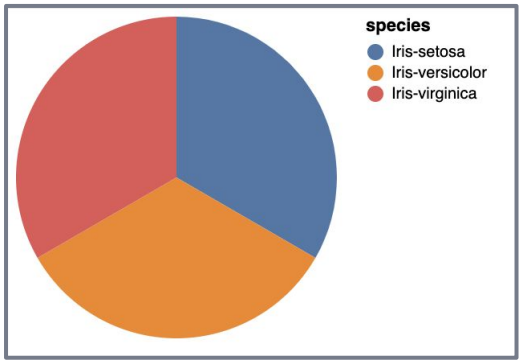
Genetic representation of DBo

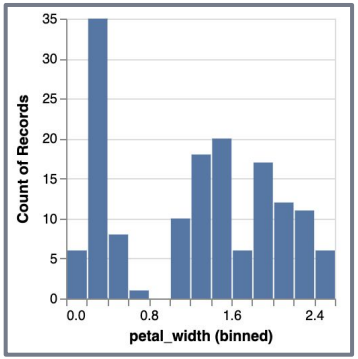
Genes and Chromosomes

- **Dashboard** (Individual/Chromosomes)
- **Visualization** (big genes)
- **Data Attributes** (small genes)
- **Visual Attributes** (small genes)

Vis.	sepal_length	sepal_width	petal_length	petal_width	species
scatterplot		x		y	color
Piechart					color
Histogram				x	







Crossover

Parent (a)

Vis.	A1 (num)	A2 (num)	A3 (nomi)	A4 (nomi)
scatterplot	x	y	color	
Barchart				color

Cutting Point

Parent (d)

Vis.	A1 (num)	A2 (num)	A3 (nomi)	A4 (nomi)
scatterplot	x	y		color
Piechart				color
Barchart		x		color
Histogram	x			
Barchart			color	

After Crossover

offspring_1

Vis.	A1 (num)	A2 (num)	A3 (nomi)	A4 (nomi)
scatterplot	x	y	color	
Piechart				color
Barchart		x		color
Histogram	x			
Barchart			color	

offspring_2

Vis.	A1 (num)	A2 (num)	A3 (nomi)	A4 (nomi)
scatterplot	x	y		color
Barchart				color

Mutation

Randomly selected offspring: offspring_1

New DBo (1): Mapping (Data Attribute ← → Visual Attributes)

Vis.	A1 (num)	A2 (num)	A3 (nomi)	A4 (nomi)
scatterplot	x	y		color
Piechart				color
Barchart	x			color
Histogram	x			
Barchart	x			color

New DBo (2): Added Visualization

Vis.	A1 (num)	A2 (num)	A3 (nomi)	A4 (nomi)
scatterplot	x	y	color	
Piechart				color
Barchart		x		color
Histogram	x			
Barchart			color	
Histogram		x		

New DBo (3): Removed Visualization

Vis.	A1 (num)	A2 (num)	A3 (nomi)	A4 (nomi)
scatterplot	x	y	color	
Barchart				color
Histogram	x			
Barchart			color	

Genetic algorithm (GA) for dashboard generation

DBo Evaluation function

Data Attribute Name	sepal_length	sepal_width	petal_length	petal_width	species
Unique Values	35	23	43	22	3
Data Attribute Type	Numeric	Numeric	Numeric	Numeric	Nominal
User Importance	75	50	50	25	100

$$F(D) = \alpha\ Match + \beta\ Cov_{imp} - \gamma\ Cplx$$

Match = Match Score

(Mapping of Visual Attribute with Data Attributes using Mackinlay Matrix)

= Sum of matching quality values (using Mackinlay Matrix) / no. of VA in vis. * (100)

Cov_{imp} = Coverage Importance

(Data attributes covered)

= Sum of DA importance / total no. DA in dataset * (100)

Cplx = Complexity

(ease of understanding the visualization)

= Sum of all vis. complexity present in DBo/ no. of vis in DBo * (100)

Cof./U	NV	BE	EX
α	1	1	1
β	2	2	2
γ	1	0.5	0.1

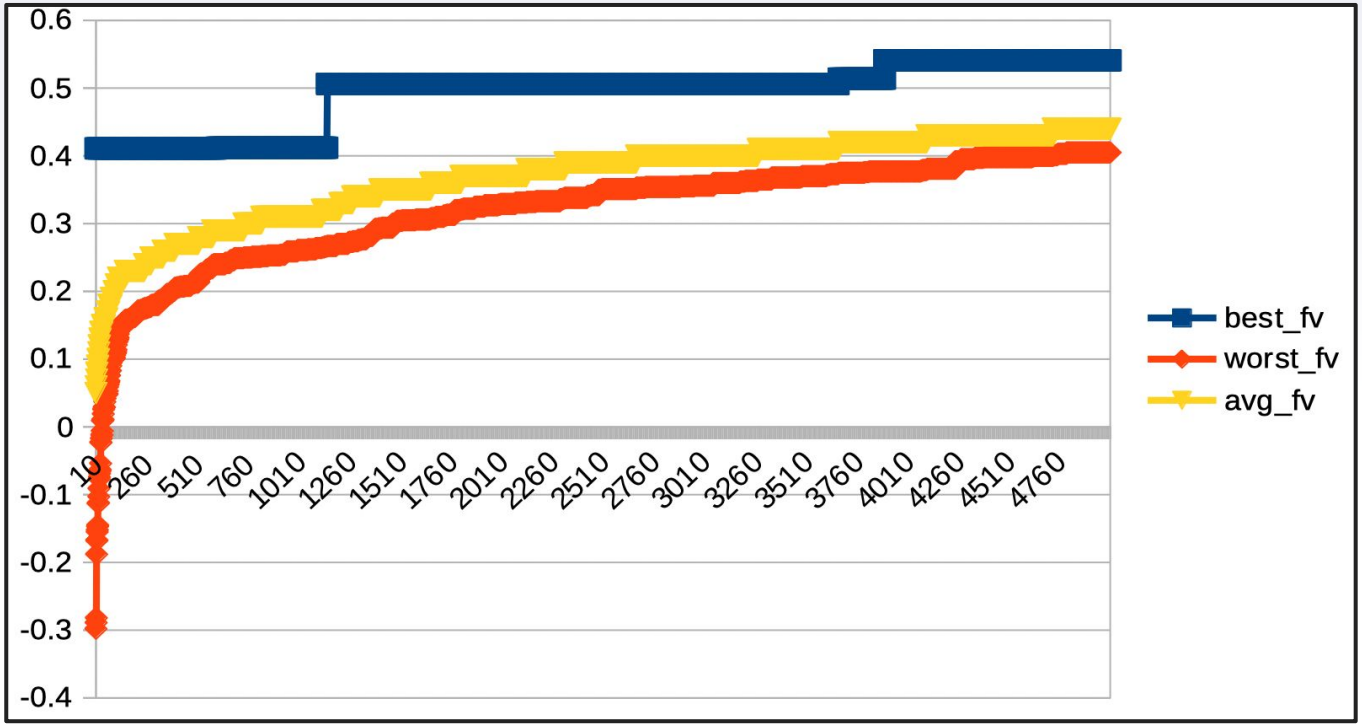
Grid Search: outcomes for Scenario 3

(to study GA parameters)

Generation / pop. size	10	20	30	50	100
1K	0.467 (0.057)	0.527 (0.056)	0.520 (0.066)	0.483 (0.052)	0.508 (0.016)
2K	0.444 (0.043)	0.581 (0.047)	0.563 (0.055)	0.573 (0.062)	0.554 (0.041)
5K	0.609 (0.057)	0.594 (0.053)	0.607 (0.041)	0.608 (0.043)	0.638 (0.039)
10K	0.618 (0.067)	0.639 (0.049)	0.638 (0.022)	0.640 (0.019)	0.647 (0.022)
20K	0.623 (0.038)	0.655 (0.038)	0.633 (0.031)	**0.699 (0.027)**	0.681 (0.033)

Syntex: **Average fitness** over **10 runs** of the **best individual** found at **each run** (**Standard Deviation**)

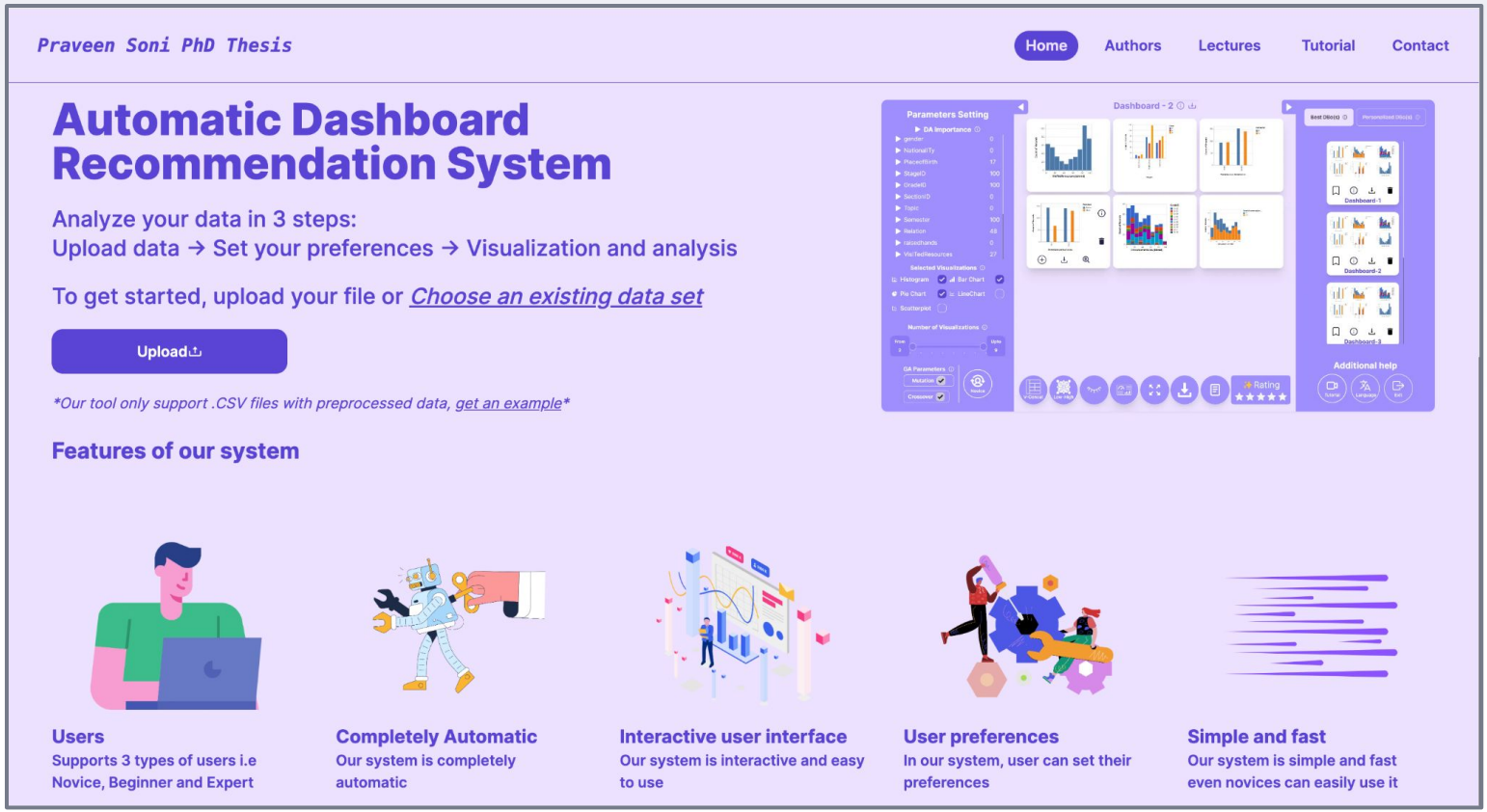
Evaluation function run time



Fitness function is fast to compute (14K fitness evaluations per second)

A run for 20K generation with 50 DBo as population (6.4s)

SimpleVis tool and its features - Demo



https://simplevis.soniconsultancy.in/

System Design

System requirements and goals:

Novices	Common users	Experts
- Simple and fewer efforts, - Completely Automatic	- Parameter setting	
- Simple terminologies and Interactive user interface - DBos customized and Personalized DBo creation - Visualization manual editing		- GA parameters and Fitness function manual editing
- Fast and quick recommendation - 5 Best DBos Recommendations - Dashboard Rating (Feedback and suggestions) - Optional Manual editing		

Participants and Datasets

Parameter	Details
Participants	25 (18M + 7F)
Age group	20-29 (21.48 Mean and 2.45 SD)
Level of Study	18 Bachelors, 2 Masters, 1 Doctorate, 4 Corporate Employees
Level of BI Tool Mastery	11 Novices (new), 14 participants (familiar)
Datasets	**Training:** Iris **Task 1:** Car, Weather (with same no. of attributes and its types) **Task 2:** Wine, Crop (with same no. of attributes and its types)

Evaluation Protocol

Parameter	Task-Tool	Steps
Exploratory Task (DBo with all data attributes)	T1-PBI	1. Welcome and general introduction 2. Fill participant basic details 3. Training 4. T1 with SV and PBI 5. T2 with SV and PBI 6. Feedback and questionnaire form
	T1-SV	
Directed Task (DBo with at least 2 scatterplot)	T2-PBI	
	T2-SV	

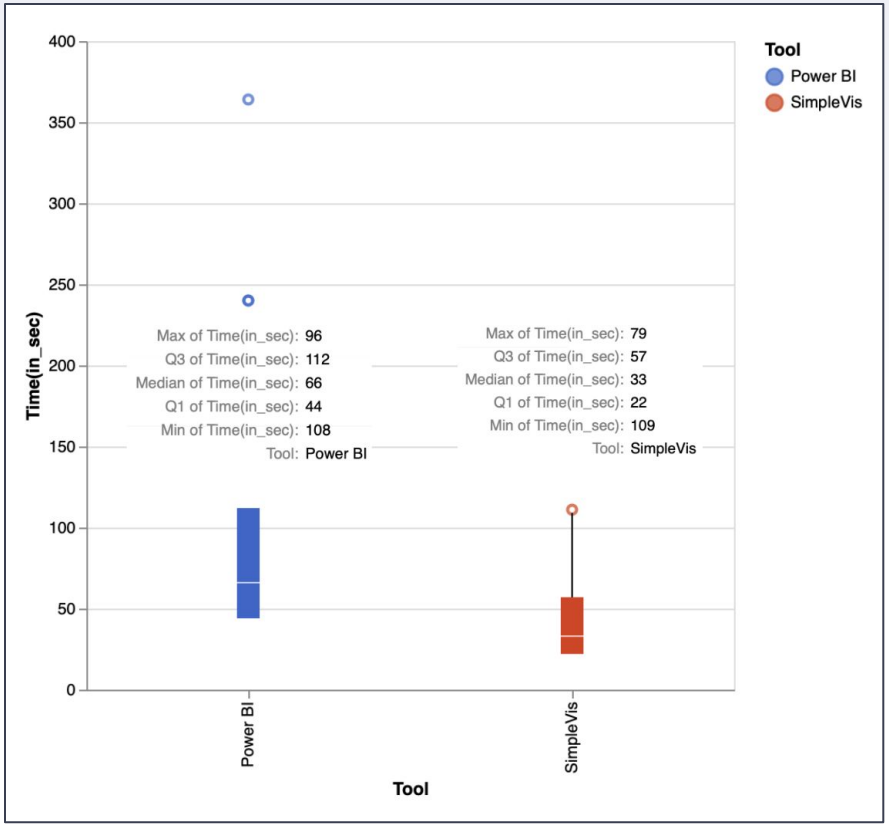
Total time: **72 mins** (5+ 5+ 28 + 12 + 12 + 10)

Task based: Time and Quality

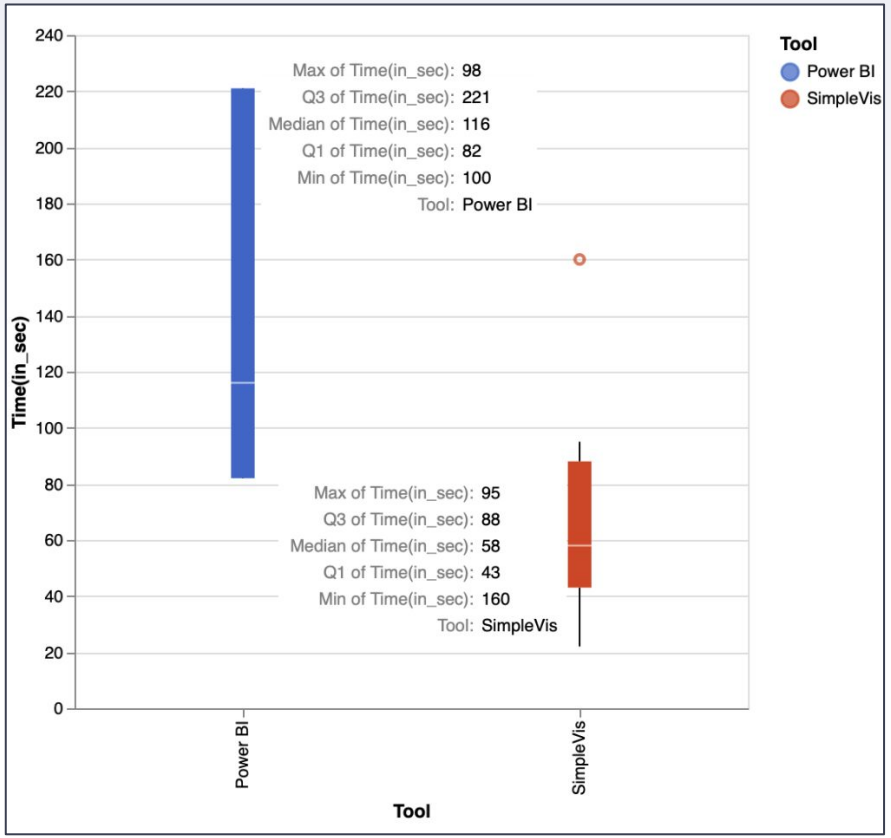
Task	*Dimension*	*TaskTool*	Mean	SD	p-value	Remark
Task 1	Time	T1-PBI	101.48	85.40	0.002	SV faster than PBI
		T1-SV	42.72	27.14		
	Quality	T1-PBI	4.48	0.37	0.85	SV similar quality as PBI
		T1-SV	4.80	1		
Task 2	Time	T2-PBI	153.6	93.09	0.0001	SV faster than PBI
		T2-SV	64.4	29.51		
	Quality	T2-PBI	1.48	0.77	0.15	SV similar quality as PBI
		T2-SV	1.76	0.60		

SV is faster than PBI and obtained similar quality of results for both the tasks

Task based: Time Analysis with BoxPlot



Task-1



Task-2

User Type based: Time and Quality

Task	*Dimension*	*TaskToolUser*	Mean	SD	p-value	Remark
Task 1	Time	T1-PBI-NV	101.27	110.87	0.68	NV almost similar to EXP
		T1-PBI-EXP	94.57	62.37		
	Quality	T1-PBI-NV	4.91	0.30	0.42	NV similar quality as EXP
		T1-PBI-EXP	4.79	0.43		
	Time	T1-SV-NV	40.45	28.04	0.72	NV similar to EXP
		T1-SV-EXP	44.50	27.34		
	Quality	T1-SV-NV	5	0	0.38	NV similar quality as EXP
		T1-SV-EXP	4.46	1.34		

Task	*Dimension*	*TaskToolUser*	Mean	SD	p-value	Remark
Task 2	Time	T2-PBI-NV	104	89.89	0.01	EXP takes more than NV
		T2-PBI-EXP	192.57	77.87		
	Quality	T2-PBI-NV	1	0.89	0.003	EXP better quality than NV
		T2-PBI-EXP	1.86	0.36		
	Time	T2-SV-NV	62.73	37.60	0.80	NV similar to EXP
		T2-SV-EXP	65.71	22.71		
	Quality	T2-SV-NV	1.55	0.82	0.11	NV similar quality as EXP
		T2-SV-EXP	1.93	0.27		

- NV takes similar time and obtain similar quality results with both PBI and SV

- EXP takes more time and obtain better quality of results with PBI (novice give up quickly as not familiar with PBI)
- NV takes similar time and obtain similar quality of results with SV

Questionnaire Analysis

Responses/Question	Mean	SD
Q1. SimpleVis is: not useful - useful	4.4	0.76
Q2. PowerBI is: not useful - useful	3.92	0.91
Q3. Solving the tasks with SimpleVis was: Difficult - Easy	4.68	0.85
Q4. Solving the tasks with PowerBI was: Difficult - Easy	3.52	0.92
Q5. Which tool is the easiest to use? PBI - SV	4.68	0.56
Q6. Which tool is the fastest? PBI - SV	4.4	0.91

Q7. Which tool provides the most interesting initial set of visualizations? PBI - SV	3.56	1.26
Q8. Which tool facilitates the most manual editing of visualizations? PBI - SV	2.76	1.42
Q9. Which tool was the most stimulating? PBI - SV	3.76	1.13
Q10. Which system would you choose to create a dashboard? PBI - SV	3.76	1.20
Q11. Which system was the most interactive? PBI - SV	3.56	1.39

- SV is slightly more useful and easier to do task
- SV is faster and easier to use
- SV is more stimulating and interactive
- PBI support more manual editing

System Evaluation: Conclusion

1. **User-Friendliness:**
   SimpleVis is more intuitive and requires less manual effort with its automation features

2. **Speed in Complex Tasks:**
   Power BI performs slightly faster in handling complex data tasks.

Power BI

3. **User Type Supports:**
   Novices can easily learn and perform data analysis with **SimpleVis** as compared to Power BI

Conclusion

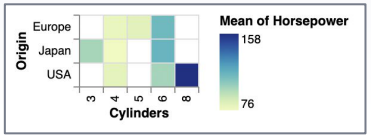
Problem Identified: Lack of **Automatic Dashboard Recommendation** for Novices, extended with explicit preference-based optimization for experts and optional manual editing.

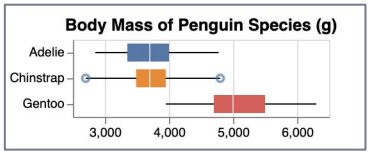
1. Systematic **literature review** on automatic dashboard generation.
2. Explicit preference-based **optimization for dashboard generation** using genetic algorithms.
3. Contribution to **dashboard automation with development** of SimpleVis Tool.
4. SimpleVis goes beyond automation by offering **support for adaptive user preferences and optional manual editing.**
5. **System validated** through user evaluations and comparative study, with its **practicality and relevance.**

System Perspective (Short Term)

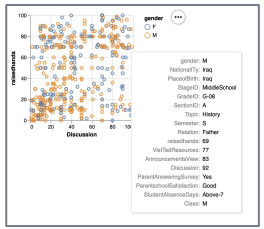
1. **Enhanced User Interface Interactivity with Automation in:**
   - ➢ **selection of data attribute type** for newly uploaded data
   - ➢ **'dashboard rating system'** suggestions implementation

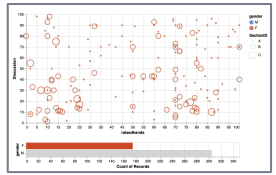
2. **Accommodating more complex datasets and diverse analytical needs with:**
   - ➢ **Extension of visualizations** by adding new visualization types





3. **Enhances analytical capabilities by empowering users to uncover more profound insights by:**
   - ➢ **Aggregation operators** (Sum, Mean, SD) integration
   - ➢ User interface effects: **Brushing, Linkage, Annotation**





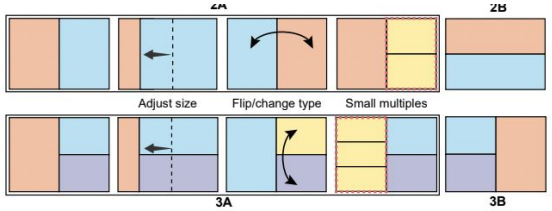
System Perspective (Long Term)

User based recommendation enhancement:

- **History** based and Project specific Recommendations
- **Predefined templates** Proposition
- Integration into third-party business tools **(API integration)**

Optimization methods and fitness functions improvements:

- **Enhancements to** Genetic Algorithm and Fitness Function
- **Global Layout** Optimization for Dashboards



Publications

Publications:

- **"A survey on automatic dashboard recommendation systems."** (Visual Informatics)
- **"A genetic algorithm for automatic dashboard generation: first results."** (In 2023 27th International Conference Information Visualisation (IV))
- **"Challenges for automatic dashboard generation systems in the context of novice users."** (In 18e journées Business Intelligence & Big Data (EDA 2022))
- **"Un outil de génération automatique de tableaux de bord."** (Atelier «La place des usagères et usagers dans les outils de fouille et d'exploration de données»(PAUL@ EGC 2024) (2023))
- **System paper to be submit in Journal (under preparation)**