# CrimeSeen: An Interactive Visualization Environment for Scenario Testing on Criminal Cocaine Networks

Frederike Oetker[1], Liza A.S. Roelofsen[1], Rob G. Belleman[1], and Rick Quax[1,2]

[1] Computational Science Lab, University of Amsterdam, The Netherlands
[2] Institute for Advanced Study, University of Amsterdam, The Netherlands
f.oetker@uva.nl

**Abstract.** The resiliency of criminal networks against law enforcement interventions has driven researchers to investigate methods of creating accurate simulated criminal networks. Despite these efforts, insights reaching law enforcement agencies remain general and insufficient, warranting a new approach. Therefore, we created CrimeSeen - an interactive visualization and simulation environment for exploring criminal network dynamics using computational models. CrimeSeen empowers law enforcement agencies with the possibility to independently test specific scenarios and identify the most effective disruption strategy before deploying it. CrimeSeen comprises of three components: Citadel, a web-based network visualization and simulation tool serving as the interface; the model, defining rules for criminal network dynamics over time, with the Criminal Cocaine Replacement Model as the use-case in this project; and the simulator, connecting the model and interface and enhancing their functionality through transformations, triggers, and statistics. CrimeSeen was evaluated with sequential usability testing, revealing a positive trend in effectiveness and efficiency over time, with mean scores exceeding 80%. However, user satisfaction did not significantly change and remained below the average for web applications, prompting recommendations for future work.
Criminal Networks Human-Computer Interaction Graph Visualization Simulation Interaction Collaboration Usability Testing Computational Modelling

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_24

## 1 Introduction

Understanding criminal cocaine networks is challenging due to their secretive nature, posing difficulties for law enforcement's intervention efforts [1, 2]. Existing approaches lack the tools necessary for effective interventions, particularly in identifying key figures like kingpins. Computational criminal network models offer promise in this regard, yet their complex nature presents accessibility challenges for law enforcement personnel without specialized training [3]. Collaboration among law enforcement agencies and stakeholders is essential for disrupting criminal networks, but integrating data from various sources is challenging due to privacy restrictions [3, 4]. CrimeSeen addresses these challenges by offering law enforcement accessible visualization and simulation tools for evaluating intervention scenarios collaboratively. By bridging the gap between researchers' models and law enforcement, CrimeSeen enables independent, efficient, and collaborative evaluation of intervention methods. Oetker et al. developed a comprehensive criminal network model of the Netherlands, but privacy laws limit data sharing, hindering the model's effectiveness [3, 4]. CrimeSeen aims to overcome these limitations by providing law enforcement with a platform to input real data and explore intervention scenarios independently. CrimeSeen, short for **CR**iminal **I**nvestigation and **M**odelling **E**nvironment for **S**cenario testing and **E**valuation using **E**xploratory interactive **N**etwork visualization, empowers law enforcement to evaluate intervention strategies effectively [3, 4]. It facilitates the visualization of different scenarios, helping domain experts identify the most effective interventions. Moreover, CrimeSeen fosters collaboration between researchers and law enforcement by providing a platform for independent and collaborative evaluation of intervention methods.

## 2 Background and Related Work

This project spans three critical research areas: criminal networks, Human-Computer Interaction (HCI), and usability testing. Criminal network research, particularly focused on drug networks in the Netherlands, is crucial for law enforcement to plan and execute interventions effectively, emphasizing the need for new disruption strategies [5, 6]. Criminal Network Theory (CNT) and Social Network Analysis (SNA) provide insights, but specific strategies must be tailored to the network structure to combat organized crime effectively [7]. In HCI, the design, evaluation, and implementation of interactive computing systems for human use are paramount, ensuring that regardless of system sophistication, usability requirements are met through iterative user-inclusive design processes [8, 9]. Usability testing, a common method for evaluating interactive systems, measures users' ability to efficiently accomplish goals, influencing their experience and future system usage [10]. To assess outcomes, metrics related to users' performance and usability experience are defined, drawing from established tasks in interactive visualization and network analysis within the criminal network context [11–14].

**Metrics** Usability is measured across three dimensions [15]: effectiveness, efficiency, and satisfaction. Effectiveness pertains to users' ability to successfully complete tasks with the system [15], [16]. Efficiency quantifies the resources expended by users to complete tasks, whether absolutely in time units or relatively as a percentage of total task time spent [15]. Both dimensions necessitate well-defined task specifications for measurement. User satisfaction with a system is assessed via standardized questionnaires

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_24

like the System Usability Scale (sus), providing a usability score between 0 and 100, requiring both qualitative and quantitative measures, encompassing issues from task completion to navigation [17–21].

**Collaborative and Interactive Visualization** Collaborative visualization, a facet of Computer-Supported Cooperative Work (cscw), enhances collaboration among participants by employing shared visualizations, crucial for addressing complex problems requiring input from multiple individuals and levels of engagement [22, 23]. Interactive systems offer greater insight and visualization possibilities compared to static visualizations, addressing challenges like cluttering and scalability in network visualization [12]. However, current network visualization applications lack specialized interventions for law enforcement, necessitating an integrated system providing direct accessibility of criminal network models to law enforcement agencies [24].

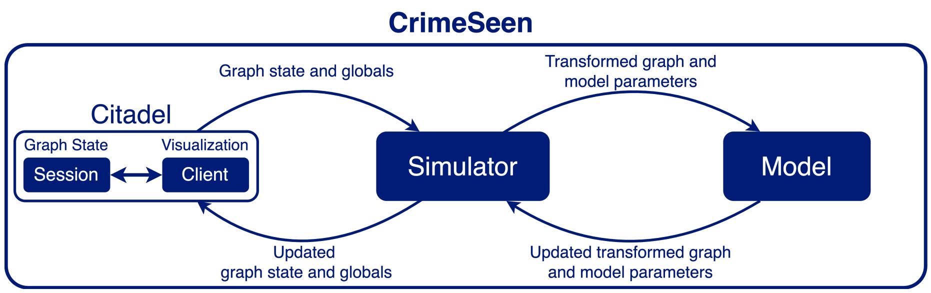


**Fig. 1.** *The three main components of CrimeSeen, including their interactions in running a simulation. Citadel is further specified in the session containing the graph state and the client containing the visualization of the graph.*

## 3 System Requirements

To attain direct accessibility to criminal network models, the created system has to meet several requirements. These requirements are further specified by domain experts through semi-structured interviews. CrimeSeen aims to provide: general accessibility (easy and intuitive usage, no specific skills or programming expertise required); collaboration (facilitating collaboration among multiple parties, viewing, and sharing information); network analysis (support for all analysis tasks, including SNA techniques for criminal networks, and basic graph analysis tasks); and high level of usability (ensuring adoption by law enforcement, enjoyable user experience, intuitive access to all functionalities).

## 4 CrimeSeen

CrimeSeen consists of three key components: the interface, the model, and the simulator (Figure 1). The interface relies on Citadel, enabling easy model interaction without

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_24

coding and enhancing accessibility for law enforcement [25]. Adaptations to Citadel, such as global variables, improve analytical capabilities and model integration [25]. The simulator facilitates compatibility with Citadel, using the ccrm as a demonstration, with a stateless step function approach ensuring system integrity and adaptability [26]. The model, exemplified by the ccrm, is crucial for scenario testing and must adhere to specific conditions for proper execution within CrimeSeen [27, 26].

**Usage Scenario** In CrimeSeen, law enforcement can execute intervention strategies tailored to the Criminal Cocaine Replacement Model (CCRM), aimed at disrupting criminal cocaine networks effectively [26]. Users initiate sessions and connect to the simulator, selecting either pre-existing networks or custom graphs compatible with the simulator and model requirements [26]. Once sessions are established, users explore intervention scenarios, guided by instructions within CrimeSeen [26]. The CCRM offers intervention methods such as automated or manual kingpin removal, targeting central criminals based on centrality metrics like in-degree, out-degree, betweenness, or closeness centrality [26]. After kingpin removal, CrimeSeen allows users to influence new kingpin selection by designating an agent as a "Kingpin Candidate," potentially proposing replacements to conclave members [26]. Through experimenting with these strategies, law enforcement assesses their effectiveness in disrupting criminal networks and identifies key nodes for removal [26].

## 5 Evaluation

We conducted three sequential usability testing sessions for CrimeSeen to gather feedback and enhance the tool. These sessions, supervised and conducted in a controlled environment, lasted about an hour each, with participants using personal laptops for familiarity. Participants from diverse backgrounds were recruited to evaluate accessibility, while tasks were designed to assess CrimeSeen's analysis capabilities and overall usability [28].

**Experimental Setup** Participants engaged in an experimental setup involving a Qualtrics questionnaire with detailed instructions and sections covering network tasks, simulation running, demographics, and usability assessment. The questionnaire aimed to achieve specific goals, such as testing network understanding, gathering feedback on simulation dynamics, and evaluating usability through the SUS and open-ended feedback [15].

**Outcome Metrics** To evaluate usability, the metrics of interest are: effectiveness, efficiency and satisfaction [15]. Effectiveness was measured as the percentage of the questions answered correctly. The formula used for the computation of the percentage score for the effectiveness of a participant was:

$$Effectiveness(j) = \frac{1}{N} \sum_{i=1}^{n} \frac{SC(i,j)}{l} \cdot 100 \tag{1}$$

where $i$ is the task index, $j$ the index of the sessions participant, $n$ the total number of sub-questions, $l$ the number of sub-questions the task has and N the number of

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_24

main tasks participants were given. Since we did not define a desired speed for task completion time, we use relative efficiency, which is measured per participant using the following formula:

$$RelativeEfficiency(j) = \frac{\sum_{i=1}^{N} n_{ij} t_{ij}}{\sum_{i=1}^{N} t_{ij}} \tag{2}$$

where $j$ is the index of the participant, $i$ is the task index, $N$ the number of tasks (11), $n_{ij}$ the score of task $i$ by participant $j$ and $t_{ij}$ the time spent on task $i$ by participant $j$.

Satisfaction was assessed using the System Usability Scale (sus [15, 18]. To tailor the sus for our evaluation, minor adjustments were made to the questionnaire, detailed in A.9. The SUS score is determined by the sum of the answer scores from 0 to 4 where for the odd questions the score increases the more the user agreed with it and for even questions the score increases the more the user disagreed with it.

$$SUSscore = \sum_{i=1}^{10} sc(i) \cdot 2.5 \tag{3}$$

where:

$$sc(i) = \begin{cases} q_{score} - 1 \; if i \bmod 2 \neq 0 \\ 5 - q_{score} \; if i \bmod 2 = 0 \end{cases} \tag{4}$$

Here, $q_{score}$ is the score for the question taken from the response's position on the Likert-scale from left (1) to right (5).

## 6 Results

The analysis of the results of the usability testing involving 32 participants focuses on the development of outcome metrics between test sessions, and the impact of collaboration and interaction test settings on these metrics.

**Usability** Results from Table 1 show effectiveness, relative efficiency, and satisfaction evolution across test sessions. Effectiveness notably improves from below 50% to over 80%, with reduced score range. Relative efficiency improves between first and second sessions but less between second and third. Satisfaction scores vary across sessions without clear trend. Kruskal-Wallis H test confirms effectiveness and relative efficiency increases but not satisfaction. Kendall's test reveals strong relationships, especially between effectiveness and relative efficiency. Total time spent correlates strongly with task time and outcome metrics, increasing across sessions [28].

**Task Analysis** Analyzing task performance across sessions, we focused on completion time and effectiveness (Figure 3) [28]. Task completion times varied significantly, reflecting differences in complexity (for an overview of the tasks, refer to A.8 in Supplementary Materials). For instance, task N4 displayed decreased completion time across sessions, indicating improved efficiency likely due to enhanced familiarity and improved system usability. Concurrently, task effectiveness demonstrated a positive trend, with

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_24

**Table 1.** *The result of the Kruskal-Wallis H test for comparing the effectiveness, relative efficiency and satisfaction value per testing session. The H-statistics are shown including their p-value between brackets and significance indicated with* $^{*}\alpha < 0.05$, $^{**}\alpha < 0.01$, $^{***}\alpha < 0.001$.

| | All Sessions | Session 1 and 2 | Session 1 and 3 | Session 2 and 3 |
|---|---|---|---|---|
| **Effectiveness** | 20.7 (0.0)*** | 9.61 (0.002)** | 15.66 (0.000)*** | 6.29 (0.012)* |
| **Relative Efficiency** | 16.91 (0.000)*** | 11.57 (0.001)*** | 13.15 (0.000)*** | 0.73 (0.391) |
| **Satisfaction** | 3.6 (0.166) | 3.31 (0.069) | 1.58 (0.209) | 0.58 (0.448) |


**Fig. 2.** *The time spent in minutes on each task across sessions.*

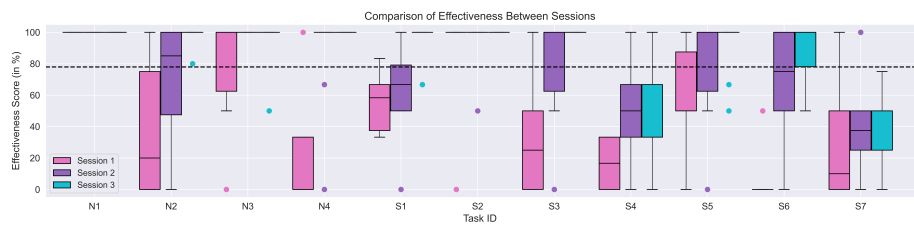


**Fig. 3.** *The effectiveness percentages for each task across sessions. The dotted line represents a threshold for comparison with the industry standard which has a task completion rate of 78%.*

an increasing percentage of participants completing tasks correctly over sessions, as evident in Figure 3. While some tasks, like N1 and S2, consistently showed high completion rates, others, like S7, remained below the industry standard even in the final session.

**Collaboration and Interaction** Figure 2 shows lower outcome metric scores in the "with collaboration" setting compared to "without collaboration". Weak negative correlations between group participation and efficiency, satisfaction, and effectiveness were observed. Participants in groups tended to spend moderately more time on tasks and questionnaires, with no significant differences in medians between groups [28].

## 7 Discussion

Sequential usability testing showed improvements in effectiveness and efficiency across sessions, while user satisfaction remained relatively stable (Figure 2). Task completion

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_24

**Table 2.** *The result of the Kruskal-Wallis H test for comparing the effectiveness, relative efficiency and satisfaction value for the collaboration and interaction test setting. The H-statistics are shown including their p-value between brackets and if applicable, significance is indicated with* $^*\alpha < 0.05$, $^{**}\alpha < 0.01$, $^{***}\alpha < 0.001$.

| | **Collaboration** | **Interaction** |
|---|---|---|
| **Effectiveness** | 0.13 (0.718) | 0.72 (0.396) |
| **Relative Efficiency** | 0.69 (0.403) | 0.01 (0.940) |
| **Satisfaction** | 1.62 (0.203) | 0.39 (0.533) |

rates were generally high, though some tasks proved challenging, indicating potential overestimation of difficulty. Usability issues persisted, highlighting CrimeSeen's ongoing development stage and the need for further enhancements. Participant demographics had minimal impact on outcomes, suggesting consistency in CrimeSeen's performance across users and devices (for an overview of the participants' demographics, see Figure S3 in the Supplementary Materials).

**Collaboration and Interaction** The results from both collaborative and interaction settings showed no significant differences among specific test groups, as evident in 2. This unexpected finding may be due to small group sizes and the distribution of these settings across sessions, limiting the detection of significant effects. Additionally, the implementation of these settings may have influenced the outcomes, with participants primarily working individually despite the collaborative setting, leading to challenges in task coordination and differing time perspectives.

**Limitations** The interaction setting solely involved engagement with the simulation, providing participants with similar test conditions until the simulation phase. This minimal influence may explain the lack of significant results, influenced by resource constraints and the absence of law enforcement end users. However, diverse participants effectively utilized the system.

## 8 Conclusion

CrimeSeen integrates the CCRM model into Citadel, providing law enforcement with a comprehensive network visualization and simulation tool. Through iterative evaluations, we identified usability issues and desired features, informing future developments focused on scalability, security, and statistical significance. As a prototype, CrimeSeen requires further development before deployment in law enforcement, with ongoing efforts aimed at refining and enhancing its capabilities. Ultimately, CrimeSeen enables independent simulation of disruption strategies and optimizes resource allocation for more effective interventions.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_24

## References


1. E. R. Kleemans, Organized Crime and the Visible Hand: A Theoretical Critique on the Economic Analysis of Organized Crime, Criminology & Criminal Justice 13 (5) (2013) 615–629. doi:10.1177/1748895812465296.
2. P. J. Carrington, The SAGE Handbook of Social Network Analysis, Sage Thousand Oaks, CA, 2011, Ch. Crime and Social Network Analysis, pp. 236–255. doi:10.4135/9781446294413.
3. F. Oetker, V. Nespeca, T. Vis, P. Duijn, R. Quax, Framework for Developing Quantitative Agent-Based Models Based on Qualitative Expert Knowledge: An Organised Crime Use Case, [Unpublished manuscript] (2023).
4. T. Diviák, Key Aspects of Covert Networks Data Collection: Problems, Challenges, and Opportunities, Social Networks 69 (2022) 160–169. doi:10.1016/j.socnet.2019.10.002.
5. J. Preece, Y. Rogers, H. Sharp, D. Benyon, S. Holland, T. Carey, Human-Computer Interaction, Addison-Wesley Longman Ltd., 1994.
6. G. Sinha, R. Shahi, M. Shankar, Human Computer Interaction, in: 2010 3rd International Conference on Emerging Trends in Engineering and Technology, 2010, pp. 1–4. doi:10.1109/ICETET.2010.85.
7. F. Karray, M. Alemzadeh, J. A. Saleh, M. N. Arab, Human-Computer Interaction: Overview on State of the Art, International Journal on Smart Sensing and Intelligent Systems 1 (1) (2008) 137–159. doi:10.21307/ijssis-2017-283.
8. T. Issa, P. Isaias, Sustainable Design: HCI, Usability and Environmental Concerns, 2nd Edition, Springer, 2022.
9. J. S. Dumas, J. E. Fox, Usability testing: Current practice and future directions, in: Human-Computer Interaction, CRC Press, 2009, pp. 247–268.
10. J. Niès, S. Pelayo, From Users Involvement to Users' Needs Understanding: A Case Study, International journal of medical informatics 79 (4) (2010) 76–82. doi:10.1016/j.ijmedinf.2009.06.007.
11. J. J. Xu, H. Chen, CrimeNet Explorer: a Framework for Criminal Network Knowledge Discovery, ACM Transactions on Information Systems 23 (2) (2005) 201–226. doi:10.1145/1059981.1059984.
12. A. Rasheed, U. K. Wiil, M. Niazi, Evaluating PEVNET: A Framework for Visualization of Criminal Networks, in: L. Wang, S. Uesugi, I.-H. Ting, K. Okuhara, K. Wang (Eds.), Multidisciplinary Social Networks Research, Vol. 540, MISNC, Springer, Berlin, Heidelberg, 2015, pp. 131–149. doi:10.1007/978-3-662-48319-0$_1$1. $URLhttps : //link.springer.com/chapter/10.1007/978 - 3 - 662 - 48319 - 0_1 1$
13. M. K. Sparrow, The Application of Network Analysis to Criminal Intelligence: An Assessment of the Prospects, Social Networks 13 (3) (1991) 251–274. doi:10.1016/0378-8733(91)90008-H.
14. D. McAndrew, The Structural Analysis of Criminal Networks, in: The Social Psychology of Crime, Routledge, 2021, pp. 51–94.
15. J. Brooke, SUS: A Quick and Dirty Usability Scale, Usability Evaluation in Industry 189 (194) (1996) 4–7.
16. J. Sauro, What Is A Good Task-Completion Rate? (2011, March 21). URL https://measuringu.com/task-completion/
17. T. Tullis, J. Stetson, A Comparison of Questionnaires for Assessing Website Usability, in: Proceedings of the UsabilityProfessionals Association (UPA), Minneapolis, MN, USA, 7–11 June 2004, ResearchGate, 2004.



ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_24

18. S. Peres, T. Pham, R. Phillips, Validation of the System Usability Scale (SUS), in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 57, SAGE Publications, 2013, pp. 192–196. doi:10.1177/1541931213571043.
19. A. Bangor, P. T. Kortum, J. T. Miller, An Empirical Evaluation of the System Usability Scale, International Journal of Human-Computer Interaction 24 (2008) 574–594. doi:10.1080/10447310802205776.
20. G. Ellis, A. Dix, An Explorative Analysis of User Evaluation Studies in Information Visualisation, in: E. Bertini, C. Plaisant, G. Santucci (Eds.), Proceedings of the 2006 AVI Workshop on BEyond time and errors: novel evaluation methods for information visualization, BELIV, AMC Press, 2006, pp. 1–7. doi:10.1145/1168149.1168152.
21. T. Tullis, B. Albert, Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics, Interactive Technologies, Morgan Kaufmann, San Francisco, 2008. doi:10.1016/B978-0-12-373558-4.X0001-5.
22. P. Isenberg, N. Elmqvist, J. Scholtz, D. Cernea, K.-L. Ma, H. Hagen, Collaborative Visualization: Definition, Challenges, and Research Agenda, Information Visualization 10 (4) (2011) 310–326. doi:10.1177/1473871611412817.
URL https://inria.hal.science/inria-00638540
23. W. A. Wulf, The National Collaboratory - A White Paper, in: J. Lederberg, K. Uncaphar (Eds.), Towards a National Collaboratory: Report of an Invitational Workshop at the Rockefeller University, Rockefeller University, Washington, D.C., 1989, pp. 17–18.
24. J. Bertin, Sémiologie Graphique: Les diagrammes, les réseaux, les cartes, Gauthier-Villars, Paris, 1967, (Translation 1983. Semiology of Graphics by William J. Berg).
25. L. A. S. Roelofsen, F. Oetker, M. Van der Lely, R. G. Belleman, R. Quax, CitadelPolice: An Interactive Visualization Environment for Scenario Testing on Criminal Networks, in: EuroVis 2023 - Posters, The Eurographics Association, 2023, pp. 17–19. doi:10.2312/evp.20231057.
26. F. Oetker, V. Nespeca, T. Vis, P. Duijn, P. Sloot, R. Quax, Framework for developing quantitative agent-based models based on qualitative expert knowledge: an organised crime use-case (2023). arXiv:2308.00505.
URL https://arxiv.org/abs/2308.00505
27. J. Kazil, D. Masad, A. Crooks, Utilizing Python for Agent-Based Modeling: The Mesa Framework, in: R. Thomson, H. Bisgin, C. Dancy, A. Hyder, M. Hussain (Eds.), Social, Cultural, and Behavioral Modeling, Vol. 12268, Springer International Publishing, 2020, pp. 308–317. doi:10.1007/978-3-030-61255-$9_3$0.
28. T. Issa, P. Isaias, Usability and Human–Computer Interaction (HCI), in: T. Issa, P. Isaias (Eds.), Sustainable Design: HCI, Usability and Environmental Concerns, Springer, London, 2022, pp. 23–40. doi:10.1007/978-1-4471-7513-$1_2$.
29. M. Van Der Lely, A Novel Visualisation Paradigm Using AR Superimposition, Bachelor Thesis, University of Amsterdam (July 2022).
URL https://scripties.uba.uva.nl/search?id=728606
30. A. Bangor, P. Kortum, J. Miller, Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale, Journal of usability studies 4 (3) (2009) 114–123.
31. B. Shneiderman, The Eyes Have It: a Task by Data Type Taxonomy for Information Visualizations, in: Proceedings 1996 IEEE Symposium on Visual Languages, 1996, pp. 336–343. doi:10.1109/VL.1996.545307.
32. W. Kenton, Value Networks: Definition, Benefits and Types (December 2022).
URL https://www.investopedia.com/terms/v/value-network.asp



ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63759-9_24