

# Can Inverse Transparency Enable Sensible People Analytics in the Real World?

Master's Thesis

**Supervisor:** Prof. Dr. Alexander Pretschner  
**Advisor:** Valentin Zieglmeier  
**Email:** {pretschn, zieglmev}@in.tum.de  
**Phone:** +49 (89) 289 - 17834  
**Starting date:** immediately

Informatik 4 – Lehrstuhl für Software und Systems Engineering  
 Prof. Dr. Alexander Pretschner  
 Fakultät für Informatik  
 Technische Universität München

Boltzmannstraße 3  
 85748 Garching bei München

<https://www.in.tum.de/i04/>

## Context

Nowadays, employees perform their tasks often with the help of software tools that store their personal data. These data are required for the operation of the business or to improve productivity. Examples for such tools are Issue Tracking Systems (ITS), Peer Code Review (PCR) or Version Control Systems (VCS). Data they store includes the type and number of tasks completed, comments posted, or code changes. All of these data can not only be used for business processes, but they also offer opportunities for profiling of employees. To protect employees, any usage beyond business purposes for these stored data is often forbidden.

The concept of “Inverse Transparency” tries to improve upon the existing protections by giving data owners more sovereignty in how their data are used. Its core idea is to enable access to data on a more case-by-case basis, but to monitor all accesses and make those visible to data owners. On the one hand, this can help to raise awareness of data usages and better protect the employee’s personal data, on the other hand it may enable usages of data useful to teams and individuals alike.

Employee’s personal data may, for example, be useful as a more objective measure for their performance, compared to the common practice of personal judgment by their superiors. Furthermore, teams might profit from insights gained from natural language analysis or automated progress evaluation. Examples for literature on usages for these data can be found for ITS [1–3], PCR [4–6], and VCS [7–8].

## Goal

The goal of this thesis is to collect real-world data to evaluate these ideas. Therefore, it is designed as a case study in cooperation with a large German media company (~ 2000 employees, > € 1 billion yearly revenue; from here on referred to as “company S”). Firstly, the status quo at company S with regard to people analytics is to be assessed, and potential uses for employee data are to be derived from literature. Secondly, analysis plugins are to be implemented with Inverse Transparency “built in” and deployed at company S. Finally, the interest in these uses for data incorporating the added transparency is to be evaluated for data consumers and data owners. The focus lies on the widely used ITS Jira Software<sup>1</sup>, but the work should be generalizable to other tools for issue tracking.

Work is split into three phases: The theoretical, the practical, and the evaluation phase.

In the theoretical phase, the status quo at company S, as well as the literature on ITS mining for people analytics is analyzed. The status quo at company S is analyzed to determine, on the one hand, the legal frameworks limiting the use of personal data, including laws and company agreements. Also, the level of utilization of data for employee assessment, team leading, and managing is determined. On the other hand, the level of awareness and agreement of employees about their data being used to evaluate their performance is assessed. These initial interviews can help collect potential use cases for data.

In parallel, the literature on ITS mining for performance assessment is surveyed in a semi-systematic fashion. Examples for works in this field are listed in the “Context” section. The goal is to derive actionable concepts and categorize them by the data sources used, the insights generated, and their applicability at company S considering the status quo of laws and company agreements. Furthermore, existing evaluations of the effectiveness of these approaches (if available) are collected and discussed.

In the practical phase, the findings from the theoretical phase are applied to implement people analytics tools. Based on the potential use cases for data in employee assessment derived from the analysis phase, analytics plugins for Jira Software are developed as dashboard plugins. All plugins will integrate *Transparency by Design*. The plugins are then installed on

<sup>1</sup><https://www.atlassian.com/software/jira>

the Jira system at company S or on a test instance. This enables stakeholders to experience potential usages for personal data in practice.

In the evaluation phase, stakeholder opinions are collected with evaluation instruments such as interviews, questionnaires, or vignette studies. After presenting the implemented analytics tools, the level of interest of potential data consumers (managers, team leads, but also employees) in employing these data analytics, as well as the level of concern of data owners about these types of analytics, is assessed. Then, the concept of Inverse Transparency is introduced and explained, and perspectives on it collected. A focus for this final evaluation dimension lies on assessing whether Inverse Transparency can address potential concerns of stakeholders, or if they conversely see potential issues with the added transparency itself. The willingness to share of data owners will be a specific focus, comparing the results with those from their initial impressions.

## Work Plan

1. Assess status quo at company S to determine the legal context for usage of employees personal data, and the current level of utilization of these data for analytics.
2. Research related literature on people analytics for performance assessment, team organization, and managing.
3. Implement ITS data analytics plugins incorporating Inverse Transparency “by Design”.
4. Evaluate the implemented plugins with stakeholders at company S, collecting their impressions of the plugins as well as if Inverse Transparency has an influence on their judgment.
5. Document the work in the thesis.

## Deliverables

- Source code of the implemented dashboard plugins.
- Evaluation data, such as questionnaire responses or interview transcripts.
- Thesis written in conformance with TUM guidelines.

## References

- [1] Rastogi, Ayushi et al. “Samiksha: Mining Issue Tracking System for Contribution and Performance Assessment” *Proceedings of the 6th India Software Engineering Conference* (2013).
- [2] Gousios, Georgios et al. “Measuring Developer Contribution from Software Repository Data” *Proceedings of the 2008 international working conference on Mining software repositories* (2008).
- [3] Dane Bertram. “The Social Nature of Issue Tracking in Software Engineering.” Master’s Thesis, *University of Calgary* (2009).
- [4] Kononenko, Oleksii et al. “Investigating code review quality: Do people and participation matter?” *Proceedings of the 2015 IEEE International Conference on Software Maintenance and Evolution (ICSME)* (2015).
- [5] Kononenko, Oleksii et al. “Code review quality: how developers see it” *Proceedings of the 38th International Conference on Software Engineering* (2016).
- [6] Bosu, Amiangshu et al. “Characteristics of useful code reviews: an empirical study at Microsoft” *Proceedings of the 12th Working Conference on Mining Software Repositories* (2015).
- [7] Jotov, Vladimir “An investigation on the approaches for version control systems” *Proceedings of the 9th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing* (2008).
- [8] Jermakovics, Andrejs et al. “Mining and visualizing developer networks from version control systems” *Proceedings of the 4th International Workshop on Cooperative and Human Aspects of Software Engineering* (2011).