www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

AUTOMATIC DETECTION OF THE HEALTH STATUS OF WORKPLACES BY PROCESSING BUILDING END-USERS' MAINTENANCE REQUESTS

SUBMITTED: October 2024
REVISED: April 2025
PUBLISHED: May 2025
EDITOR: Žiga Turk
DOI: 10.36680/j.itcon.2025.027

Marco D'Orazio, Prof.
DICEA, Università Politecnica delle Marche, via brecce bianche, Ancona, Italy
ORCID: https://orcid.org/0000-0003-3779-4361
m.dorazio@staff.univpm.it

Gabriele Bernardini, Dr.
DICEA, Università Politecnica delle Marche, via brecce bianche, Ancona, Italy
ORCID: https://orcid.org/0000-0002-7381-4537
g.bernardini@staff.univpm.it

Elisa Di Giuseppe, Prof.
DICEA, Università Politecnica delle Marche, via brecce bianche, Ancona, Italy
ORCID: https://orcid.org/0000-0003-2073-1030
e.digiuseppe@staff.univpm.it

SUMMARY*: This paper addresses the challenge of assessing workplace health through building maintenance requests' data, particularly focusing on the impact of maintenance conditions on workers' satisfaction, well-being and possible stress levels. A data-driven methodology based on CMMS (Computer Management Maintenance Systems) is proposed, utilizing indexes to measure both the quantity and perceived quality of maintenance interventions. Sentiment and emotion analysis, along with lexical diversity indices, are applied to capture the perceptions of end-users and technical staff. The methodology successfully identifies maintenance issues in buildings and highlights differences in perception between workers' typologies. The results provide valuable insights for facility managers and organizations, enabling better-informed decisions on maintenance priorities based on both objective data and workers' feedback. This approach paves the way for future research integrating qualitative and quantitative data in facility management, with the potential to enhance decision-making and improve workplace health.*

KEYWORDS*: Building maintenance, sentiment and emotion analysis, natural language processing, user perception, workplace health status, data-driven approach.*

REFERENCE*: Marco D'Orazio, Gabriele Bernardini & Elisa Di Giuseppe (2025). Automatic detection of the health status of workplaces by processing building end-users' maintenance requests. Journal of Information Technology in Construction (ITcon), Vol. 30, pg. 650-678, DOI: 10.36680/j.itcon.2025.027*

1. INTRODUCTION

Facilities management combines multiple disciplines to maintain and enhance the built environment, guaranteeing the safety and quality of workplaces. On the one side, effective facility management (Campos Fialho *et al.*, 2023; Ensafi *et al.*, 2023; Rampini and Cecconi, 2022; Shalabi and Turkan, 2020; Yousefli *et al.*, 2020) is crucial to the success of an organization, contributing to the delivery of its strategic and operational objectives, avoiding long interruptions of critical services, and reducing the possibility of critical workers' safety levels (Chanter and Swallow, 2007; Sanni-Anibire *et al.*, 2021). On the one hand, it hence ensures the quality of the spaces and the best conditions for users' activity, mainly considering the one of workers (Cotts *et al.*, 2010).

The quality of workplaces refers not only to the degree of compliance with specifications or standards but also to an individual's relativistic evaluation (Carvajal-Arango *et al.*, 2021; Ensafi *et al.*, 2023; Zalejska-Jonsson, 2020). Users have their own definition of quality and use this reference frame in the evaluation process of workplaces (Baines *et al.*, 2017; Zhang *et al.*, 2024). Therefore, personal beliefs and values can affect the perception of workplace quality, also impacting workers' productivity (Dam-Krogh *et al.*, 2024; Rasheed *et al.*, 2021) and contributing to job-related stress conditions (Alberdi *et al.*, 2016a; Carvajal-Arango *et al.*, 2021; Jaafar, 2021). The physiological and cognitive performances of the users are strongly affected by the physical characteristics of the workplaces (Banbury and Berry, 2005; Latini *et al.*, 2021, 2023, 2024; Nolé Fajardo *et al.*, 2023; Peter *et al.*, 2023; Presti *et al.*, 2022; Yin *et al.*, 2018; Zhang *et al.*, 2024). Workers' concentration (Burmeister *et al.*, 2018), memory and task-switching ability (Vandierendonck *et al.*, 2010) can be also altered by personal permanent or temporary stressful conditions due to negative past experiences (Valiyappurakkal, 2022) and to the permanency/recurrency of issues in the working space (Bortolini and Forcada, 2018; D'Orazio *et al.*, 2022; Izobo-Martins *et al.*, 2018). Collecting continuous information about user perception of the working environment is then a crucial point to recognize in advance, reduce workplace-related stress conditions and improve productivity (Doukari *et al.*, 2023; Jahanger *et al.*, 2021; Zalejska-Jonsson, 2020).

The influence of the quality of the workplace on users' productivity and stress conditions can be analysed based on lexical features analysis (Alberdi *et al.*, 2016b; Hu and Liu, 2004a) and sentiment and emotion analysis methods (Cambria, 2016; Cambria *et al.*, 2020; Esuli and Sebastiani, 2007; Hu and Liu, 2004b; Joshi *et al.*, 2017; Nielsen, 2011; Sánchez-Rada and Iglesias, 2019; Schmidt *et al.*, 2018; Wang *et al.*, 2020) applied to "free" texts written and exchanged by the workers. Indeed, given the recent and rapid digital transformation of the buildings' construction and management sector (Errandonea *et al.*, 2020; Jahanger *et al.*, 2021; Lu *et al.*, 2020; Wong *et al.*, 2018), artificial intelligence (Pan and Zhang, 2021), and especially deep learning (Sanzana *et al.*, 2022), text mining and natural language processing (Bugalia *et al.*, 2022; Shamshiri *et al.*, 2024; Valença *et al.*, 2024) are more and more adopted (Baek *et al.*, 2021; Sánchez-Garrido *et al.*, 2023).

Job satisfaction and occupational stressors have been recently analysed on tweets using clustering methods and sentiment analysis methodologies. Brando et al. (Brando *et al.*, 2023) applied Multiple Correspondence Analysis and Hierarchical Clustering on Principal Components methods, to identify clusters of workers highlighting the influence of different levels of experience and job roles regarding stressors, satisfaction, and feelings about work and workplaces. Jura et al. (Jura *et al.*, 2022) analysed data related to the job satisfaction of professional nurses applying clustering analysis and sentiment analysis to free-text comments. Clustering analysis was used to identify different satisfaction groups, and sentiment analysis scores associated with job satisfaction.

The use of sentiment analysis to identify the user perception of unexpected stressful events has been also proposed by Gaspar et al. (Gaspar *et al.*, 2016). Zhang et al. (Zhang *et al.*, 2021) analysed the impact of work-from-home during the coronavirus disease through sentiment analysis methods, showing positive attitudes of the workers. Alruiry (Alruily, 2023) proposed an optimized deep neural network technique, based on Convolutional Neural Networks, to forecast workers' stress conditions. Li et al. (Li *et al.*, 2021) proposed to adopt an occupant-centric approach to select building performance indicators and to reduce stress conditions on workers.

Nijhawan et al. (Nijhawan *et al.*, 2022) adopted Natural Language Processing (NLP) and Machine Learning (ML) methods to detect individuals' stress. They used a BERT model to solve sentiment classification tasks. NLP methods were also used to detect mental illness (Zhang *et al.*, 2022) and depression (Liu *et al.*, 2022; Nijhawan *et al.*, 2022). Ahuja et al. (Ahuja and Banga, 2019) proposed a method to detect the mental stress of university students through ML methods, based on the perceived stress scale developed by Cohen et al. (Cohen *et al.*, 1983). Linguistic features extracted from free texts were proposed for the early detection of stress conditions (Alberdi *et* *al.*, 2016a). Text features, such as lexical diversity, content diversity, noun and verb rate, cognition, intensity and emotive word rate were proposed for the early detection of stress conditions in humans during different types of activities (Luckman *et al.*, 2020; Saleem *et al.*, 2012). Kim et al. (Kim and Kim, 2022) analysed studies that quantify human emotions in architectural environments. Peter et al. (Peter *et al.*, 2023) examined the condition of the physical work environment of lecturers in a University, collecting information about the user perception including maintenance.

Taking into account the overall building structure and its features, facility management (Chung *et al.*, 2023; D'Orazio *et al.*, 2022; Liu and Hu, 2019; Marzouk and Enaba, 2019; Zhang *et al.*, 2024) plays a critical role in shaping end-users' perception and productivity, particularly in terms of building operation and maintenance (Kim Wing *et al.*, 2016; Valença *et al.*, 2024; Zalejska-Jonsson, 2020). Bortolini et al. (Bortolini and Forcada, 2020) then proposed methods to extract the intervention priority from the users' perception of the building systems through NLP methods. The facility managers' perception of building performance assessment has been also analyzed by other works (Assaf and Srour, 2021; Bortolini and Forcada, 2018). Similarly, in literature, a consensus has been also reached on the possibility of detecting workplace-related stress conditions of the workers based on textual analysis methods. Indeed, in large organizations, it is usual to store a large amount of information on maintenance issues based on end-users' requests in Computer Management Maintenance Systems (CMMS) (Bortolini and Forcada, 2018, 2020; Chen *et al.*, 2019; Condotta and Scanagatta, 2023; Hong *et al.*, 2022; Marocco and Garofolo, 2021; Pishdad-Bozorgi *et al.*, 2018; Wong *et al.*, 2018; Yousefli *et al.*, 2020), which generally contributes to the application of data-driven approaches (Errandonea *et al.*, 2020; Gunay *et al.*, 2019).

An end-user's maintenance request generally entails a labelled free text. Labels attributed by end-users provide useful information for the technicians, such as the user's perceived urgency, the interested systems and its localization. Each maintenance request is then usually processed by technicians, to attribute the priority, the category and the needed corresponding workforce, and to perform the necessary tasks until the issue solution (Bortolini and Forcada, 2020; D'Orazio *et al.*, 2022). Besides general information for maintenance operations (Chen *et al.*, 2019; Condotta and Scanagatta, 2023; Lu *et al.*, 2020; Shalabi and Turkan, 2020; Wong *et al.*, 2018), a CMMS hence stores textual information which can be processed with NLP methods to extract linguistic features, sentiments and emotions perceived by both users and technicians, a basis for the automatic detection of workspace-related stress conditions, to improve the quality of the workplaces and the productivity of the workers.

However, the perception of the maintenance issue by technicians can be quite different from that of users. The technician's perception is often "aseptic" because expressed by a person with adequate knowledge of the real gravity and urgency of the issue and the potential risks, and not directly living in the workplace (Bortolini and Forcada, 2020; D'Orazio *et al.*, 2024; Marocco and Garofolo, 2021; Pishdad-Bozorgi *et al.*, 2018). On the contrary, the users' perception combines the issue with personal beliefs. A similar phenomenon could be retrieved in other application fields where "urgency over importance" could guide required task prioritisation and thus affect individual quality perception (Zhu *et al.*, 2018). As an output, it could be assumed that, if the perception of "urgency" by maintenance staff is significantly different from the one of end-users due to what stated above, the programmability of interventions (when work orders will be scheduled and finished) would be different from the priority on requests provided by end-users themselves. This can potentially imply disagreement in intervention timings by end-users, increasing a potential poor quality perception and thus stress conditions, since maintenance tasks are not managed as expected (Zubair *et al.*, 2024). In this sense, dynamically detecting both perceptions (by facility managers and end-users) can then be useful for the early detection of workers' stress conditions and for measuring the impact of the building status on workers (Vizer *et al.*, 2009). Early detection methods could help to better define the useful actions to improve the quality of workplaces, suggesting where and when to perform major interventions as well as their prioritization (Bortolini and Forcada, 2018; Ensafi *et al.*, 2023; Gunay *et al.*, 2019; Kim Wing *et al.*, 2016; Maslesa and Jensen, 2019; Pärn *et al.*, 2017).

Considering that the existing literature on this topic is still limited, this paper proposed an innovative data-driven workflow to assess the workplace health status by combining maintenance needs analysis with user job-related stress and quality perception connected to maintenance needs, based on information stored in a CMMS. The status of maintenance needs is assessed by the analysis of requests management and intervention process (in terms of number and resolution timing), while job-related stress connected to maintenance needs and the perception of maintenance quality is assessed by lexical diversity, sentiment and emotion analysis. A set of indexes is proposed to measure these issues, and they have been organized to assess the specific conditions related to the organization workplaces within the overall building stock. In particular, the attention is herein focused on the maintenance of building systems, since they are generally organized in wide areas and can thus affect the performance of each whole workplace. Moreover, job-related stress assessment is performed considering both organization end-users and maintenance staff technicians, to take into account their different levels of knowledge of the buildings' maintenance tasks.

In view of the above, the paper addresses the following research questions: (RQ1) Is it possible to identify critical conditions in the building stocks thanks to the combination of the number and timing of the request and of the users' perceptions?; (RQ2) Is it possible to detect the perception of the workplace quality using sentiment and emotion analysis, and to provide insights on the impact of workplace health status and related possible stress conditions of workers using lexical diversity indexes, based on information stored in CMMS? ; (RQ3): Is there any difference between the perceptions of end-users (i.e. organization workers) and of technicians on the maintenance issues, especially considering timing and priority, given their different levels of knowledge of maintenance tasks?

These questions are answered by analyzing the information stored in a CMMS of a large university building organization. The overall methodology of the present paper is introduced at the beginning of Section 2 and organized into 4 phases detailed in Sections 2.1 to 2.4. Outcomes related to each phase are then presented in the related Section, from 3.1 to 3.4. Finally, Section 4 reports the discussion on the results and the concluding remarks.

2. METHODOLOGY

The paper has been organized into the following four main steps, as summarized by the graphical framework reported in **Figure 1**.

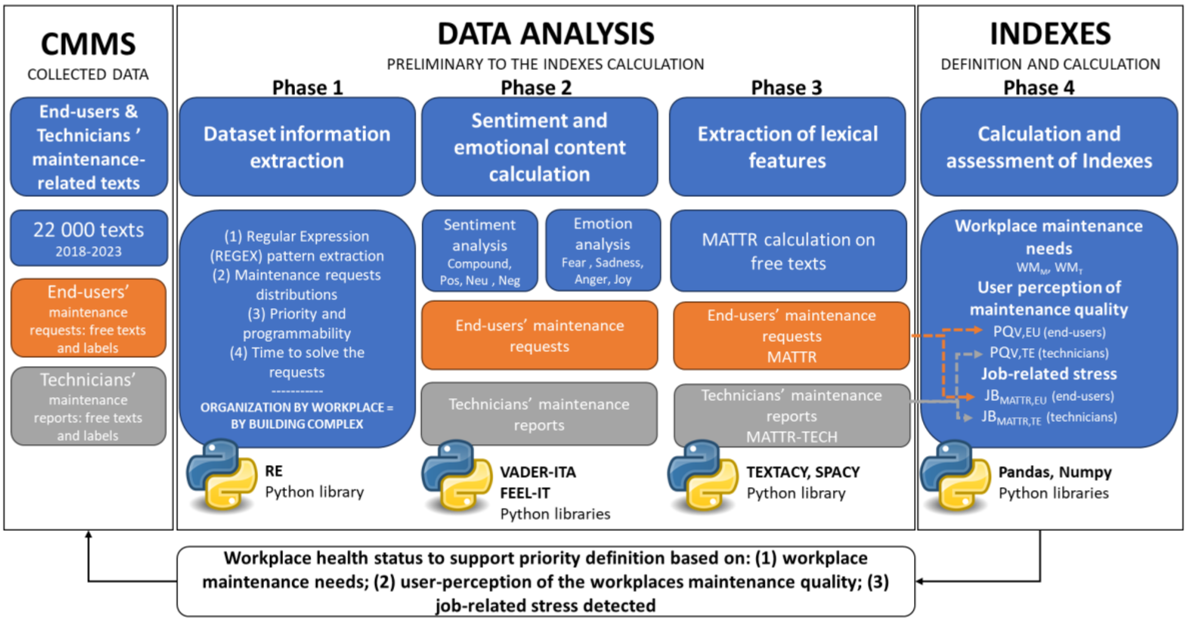

Figure 1: Research Framework.

- Phase 1 (Section 2.1): dataset information extraction and analysis, mainly oriented to the different *workplaces* in the whole organization building stock. The analysis is performed on the textual maintenance requests from the organization workers (in the following, defined as *end-users* - EU, since they are end-users of maintenance actions) and communications from maintenance staff technicians, to detect the distribution of the requests by *work category*, the assigned *priority* by workers, and *programmability* by technicians, and the effective time to reply and to complete the maintenance intervention.
- Phase 2 (Section 2.2): calculation of the sentiment (polarity) and the emotional content attributed to each communication by the organization workers and technicians, through sentiment and emotion analysis.

- Phase 3 (Section 2.3): extraction of lexical features (lexical diversity and content diversity) from the texts through NLP methods, considering both communications by workers (i.e. initial maintenance requests) and by technicians (i.e. final maintenance report after the intervention completion).
- Phase 4 (Section 2.4): definition of indexes to assess the health status and job-related stress in the workplace, with the aim of supporting the priority definition of major interventions with respect to different *workplaces*, and additional comparison between perceptions of different categories of users (i.e. workers versus technicians).

2.1 Phase 1: dataset information extraction and preliminary analysis

The dataset is composed of a corpus of more than 22,000 end-users' maintenance requests collected for more than 5 years (from January 2018 to September 2023) in a university building stock of more than 400,000 m^2. Table 1 shows the overview of the building stock, divided by the composing buildings complexes, which are associated with specific Identification codes (IDs). For each building complex, Table 1 offers an overview of the complexities characterizing the case study building stock in terms of: use, defining its "Denomination" and "Main intended use"; dimension, in terms of Gross Floor Area *GFA* [m^2]; occupancy, in terms of the approximated number of hosted workers, and the Occupancy Load (*OL*, as the ratio between the number of hosted workers and the *GFA* [pp / 100m^2]); potential obsolescence, pointing out the year(s) of construction and/or last deep renovation intervention. This kind of information is also easily available in different contexts, including those associated with BIM integration of maintenance tasks (Condotta and Scanagatta, 2023; Doukari *et al.*, 2023). Each building is generally equipped with centralized ventilation and heating/cooling systems (with fan coils).

The building stock hosts offices for more than 1300 workers (teachers, researchers and technical staff), as well as classrooms and laboratories hosting up to about 17.000 students per day. Workers can send a maintenance request to the facility managers, while students cannot directly do the same, but should ask for technical staff to report failures. As shown by Table 1, the majority of the building floor area relates to educational and research uses, and most of the workers are related to these activities.

The facility management of the whole building stock is provided by a unique external contractor. The facility management contract started in January 2018, thus the whole monitoring period refers to a single contractor. 15 technicians worked in the maintenance staff teams during the monitoring campaign, whose offices are placed in the building stocks according to the contract. Since January 2018, the contractor has collected maintenance requests coming from end-users through a CMMS. Requests have been then manipulated for analysis purposes by developing scripts in Python (rel. 3.11), using an Anaconda environment, and Dataspell 2024.2.1.1 as the IDE. "Pandas", an open-source library for data manipulation and analysis in Python, has been used to define the dataframe structure comprising the original maintenance request from the end-users, all the other information stored in the CMMS and all the analysis results about the work order defined in the following. Finally, data visualization in Python has been performed using "matlplotlib" (https://matplotlib.org/) and "seaborn" library (https://seaborn.pydata.org/).

Each maintenance request stored in the CMMS is composed of the description of the issue (text written in Italian, in a free form, without field limits), the *priority* level assigned by the end-user on a three-level scale ("*emergency"*, which is the maximum priority; "*urgent";* "*not urgent"*, that is the minimum priority) and the *date and time of the request*.

When a new request is generated by an end-user, the maintenance staff analyses the text, in order to define:

- "what" needs maintenance. According to the facility management contract, maintenance activities are organized into *work categories*, essentially comprising three main typologies related to building systems: Plumbing, Electrical, Heating and Cooling systems (including HVAC);
- "where" is placed the maintenance issue. *Work categories* rely on interventions related to building systems, and thus are generally performed at the building complex level or, even, widely affect the operation of widest areas inside each complex. Therefore, the *workplace* interested in the request is organized in the dataset by the *building complex* ID (Table 1).

These data are directly stored in the CMMS. In addition, the maintenance staff also collects additional information about each issue, e.g. by in-situ inspections. Using Python scripts, preliminary treatment for text pre-processing

(Mo *et al.*, 2020; Parisi *et al.*, 2021) has been applied to each request to reduce misspellings and remove symbols, so at to allow the application of scripts for basic automatic information extraction. According to the methodology provided by D'Orazio et al. (D'Orazio *et al.*, 2024), the following steps were taken into account. First, symbols and punctuations (e.g. "°","$", "_", "%") were removed, along with numbers. Then, misspellings (e.g. correcting typos, duplicated words) were corrected. An example of pre-processed request follows, where original misspellings and symbols are marked by strikethrough text and modifications are marked in italics: "The the following problems are reported to the bathrooms of the Villlarey buildding *building* complex of: at the groudn *ground* floor, room n.66°, anteroom accessible from room n.62°, that is to say, the study room located between the entrance of the faculty and staircases n.3°, the letf left washbasin is partially clogged and the the water drains very slowly; at the ggarage *garage*, room n.18°, bathroom block located at the end of the of the parking lot that enters from the driveway entrance remains on the left, the drain buttton button is blocked, thus water keeps flowing out of the c istern *cistern*."[1].

Once the texts have been pre-processed, Regular Expression (REGEX) pattern extraction operations have been made to detect related granular information on these issues, by using the Python (rel. 3.11) library RE in the developed scripts (https://github.com/python/cpython/tree/3.13/Lib/re/). No preliminary filtering has been used in REGEX application, allowing to trace main details and avoid misinterpreting data also in terms of contents, features and spatial location. It is worth remarking that this pre-processing has been also performed in view of lexical features extraction as reported in Section 2.2, and of sentiment and emotion analysis application defined in Section 2.3.

If the request corresponds to a real issue, the request is translated into a *work order*, and the *programmability* of interventions is attributed by the technical staff to organize the workflow and solve the issue. Besides the needed corrective actions, the *programmability* classification can also be based on the evaluation of the effective available technicians for contemporary interventions, as well as on specific materials procurement to intervene. In particular, the *programmability* of interventions is based on a four-level scale: "*not delayable*", for maximum intervention priority; "*short-term*"; "*medium-term*"; and "*long-term*" for minimum intervention priority. It can be assumed that "*not delayable*" and "*short term*" *programmability* correspond to the "*emergency*" *priority*, the "*medium-term*" *programmability* corresponds to the "*urgent*" *priority*, while "*long-term*" *programmability* corresponds to the "*non urgent*" *priority*.

After the inspection by the technicians, the work on the corrective action starts. When the issue is resolved, the time necessary to complete the work is recorded and the work order is labelled as "completed", assigning the *date and time of work closure* and providing a short report (textual communication) on the maintenance interventions (e.g. to shortly report the detected problems and the needed actions).

The time necessary to solve each request (*TIME-TO-SOLVE* [days]) has been hence calculated based on the difference between the *date and time of work closure* and on the *date and time of the request* collected into the CMMS. Nevertheless, the facility management contract also includes the definition of expected time by *work category,* depending on the *programmability* level attributed to the request. To evaluate differences among expected and effective response processes by the contractor, this work hence calculates the ratio between *TIME-TO-SOLVE* and related expected time by contract, for each request.

Given the above, the preliminary dataset analysis concerns the number of maintenance requests and work orders by *workplace* (i.e. *building complex*) and the related percentage values. The requests distribution by *work category, priority and programmability* has been also calculated in percentage terms. Statistical distributions of the ratio between *TIME-TO-SOLVE* and related expected time are also provided through Python scripts (rel. 3.11). Such an analysis allows to define the general workplace health status of the building stock, thus focusing on maintenance

[1] The original Italian text was (pre-processed issues are marked in italics): "Si segnalano le *le* seguenti problematiche bagni del complesso di *ediffici* Villlarey: al *piaano* terra, locale n.66°, antibagno accessibile dal locale n.62°, vale a dire la sala studio situata tra l'ingresso della facoltà e la scala n.3°, il lavabo di *sinsitra* è parzialmente otturato e l'l'acqua defluisce molto lentamente; nel *ggarage*, locale n.18°, blocco bagni situati nell'estremità del parcheggio che entrando dall'ingresso carrabile rimane a sinistra, il *pullsante* di scarico è bloccato, l'acqua continua a defluire dalla *c assetta*."

needs and related management issues, by buildings complex (to define differences among each of the organization workplaces) and as a whole.

*Table 1: Building stock characteristics. Each building complex is associated with an Identification code (ID), its denomination and main intended use, the related Gross Floor Area GFA, the approximated number of workers, and Occupancy Load OL, the year(s) of construction and/or the last deep renovation intervention. *: outdoor facilities, such as fields for agriculture and horticulture-related research, are excluded since they are not an object of general maintenance issues; "n.a.": not assessed in view of the specific features of the buildings complex.*

IDs	Denomination	Main intended use	GFA [m^2]	Approx. Number of workers	Occupancy Load OL [pp/ (100m^2)]	Year(s) of Construction and/or rehabilitation
L6-047-001	Rectorate and central administration	Administrative offices	5611	88	1.57	1976
L6-047-002	Central administration	Administrative offices	4678	164	3.50	1976
L6-047-003	Central administration - facilities	Garage	168	n.a.	n.a.	1976
L6-047-004	Faculty of Economics	Educational & research	59721	134	0.22	1996
L6-047-005	Faculty of Medicine and Surgery	Educational & research	65160	290	0.44	1995-2008
L6-047-006	Faculties of Engineering	Educational & research	199959	437	0.22	1990-2005
L6-047-007	Faculties of Science	Educational & research	11794	131	1.11	1997-2008
L6-047-008	Faculty of Agricultural Sciences	Educational & research	17402	93	0.53	1982-2017
L6-047-009, L6-047-010, L6-047-011, L6-047-013	Extension and Research Center in Agriculture (mainly outdoor areas + storage buildings)*	Research	3881	14	0.36	n.a.
L6-047-012	Solar Pond (outdoor facilities)	Research	0	n.a.	n.a.	n.a.
L6-047-014	Botanical garden (mainly outdoor areas + storage buildings)*	Research	1622	10	0.62	n.a.
L6-047-017	University sport facilities (indoor and outdoor)	Sport facilities	57471	n.a.	n.a.	1971

Since the facility manager is obliged to manage all the end-users' maintenance requests, it is worth noting that records with missing values were not detected. In addition, thanks to CMMS, potential missing values on the work order status were not retrieved, allowing the automatic collection of all the data about technicians' interventions associated with the starting request. A very limited number of end-users' maintenance requests still open has been detected (<0.5%). These cases have been removed from the database.

Finally, the case study application also considers the collection of data about Customer Satisfaction questionnaires administered by the Quality Assurance System Office (QASO) of the University to their workers (teachers, researchers and technical staff). Questionnaire results were collected in the same monitoring period and organized by building complex. In particular, the questionnaires include a question about the maintenance service satisfaction level according to a 6-point Likert-scale.

2.2 Phase 2: calculation of the sentiment and emotional content

Sentiment and emotion analysis was performed on textual communications pre-processed as described in Section 2.1. Nevertheless, before calculation, an additional text treatment was implemented, by excluding frequent words, but not useful for sentiment and emotion analysis (e.g. floor, room). Then, two methods were selected to extract the sentiment (polarity) and the emotional content of each maintenance request, considering both end-users' and technicians' perceptions. Sentiment and emotion analysis have been separately considered since, according to previous research, "even though these two names are sometimes used interchangeably, they differ in a few respects" (Nandwani and Verma, 2021). In particular, sentiment analysis explores positive, negative or neutral polarity of texts, determining the overall opinion expressed by the user, while emotion analysis provides deeper insights on the matter by "identifying distinct human emotion types". In addition, it is worth noting that, although emotions could have different valence (which could be theoretically similar to the concept of polarity), the comparison of outcomes from sentiment and emotion analysis would lead to relevant uncertainties and biases, especially considering the variety of emotions and how to collapse them into two or three classes. For instance, "sadness and joy are opposites, but anger is not the opposite of fear" (Nandwani and Verma, 2021), and this does not allow to simply gather emotions in opposite valences (emotion) with a direct comparison with polarities (sentiment). For this reason, this paper prefers dividing sentiment and emotion analysis, although a comparison between the related adopted methods has been provided for the sake of completeness.

From an operational standpoint, different Python scripts (rel. 3.11) were written in the developed Anaconda environment (compare Section 2.1), adopting specific libraries as in the following. Sentiment analysis was performed using VADER (valence aware dictionary for sentiment reasoning), a rule-based model for general sentiment analysis (Borg and Boldt, 2020; Hutto and Gilbert, 2014; Valença *et al.*, 2024). The effectiveness of the model has been compared to eleven benchmarks including LIWC, ANEW, General Inquirer, SentiWordNet and ML techniques (Hutto and Gilbert, 2014). VADER is based on a valence-based human-curated gold standard sentiment lexicon, sensitive to both the polarity and the intensity of sentiments expressed. The lexicon includes over 7500 lexical features with validated valence scores that indicate both the sentiment polarity and the sentiment intensity. The output of VADER algorithm for each maintenance request is a 4-tuple of scores corresponding to 3 classes of sentiments (negative, neutral, positive) and their aggregate value (compound score). VADER was initially developed for the analysis of phrases written using the English language, but translations have been made to use it with different languages, including Italian (Martinis *et al.*, 2022). VADER Italian lexicon developed by Russo et al. (Russo and Coco, 2023) has been then used to calculate positive, neutral, negative and compound scores for each sentence. A Python script was written to process the dataset with the "VADER-Italian-sentiment" Python library (https://github.com/AndreaRussoAgid/VADER-Italian-Sentiment). Such outcomes were organized by *priority* and *programmability* labels, to match sentiment contents with, respectively, end-users and technicians' assumed urgency of the interventions outlined in Section 2.1. Boxplot representations are used to provide a clear overview of sentiment analysis results distribution by focusing on the VADER compound score (*VCS*), since it evaluates both the polarity (negative, positive) and intensity of sentiment within the range from -1 (most extreme negative, included) to +1 (most extreme positive, included). The lower the VCS, the higher the perceived severity of the maintenance issue.

To extract the perceived emotions of the end-users, an emotion analysis was performed using FEEL-IT. FEEL-IT is a benchmark corpus of Italian Twitter (Bianchi *et al.*, 2021) annotated with four basic emotions ("anger", "fear", "sadness", "joy") based on the Ekman work (Ekman, 1999). FEEL-IT can predict emotions and sentiments (collapsing the emotions). FEEL-IT has been tested considering the Italian BERT model (UMBERTO) and SENTICPOL16 to evaluate the performance of sentiment and emotion classification models trained on FEEL-IT. A Python script was written to process the dataset with the "FEEL-IT" Python library.

As stated above, although VADER and FEEL-IT results are not fully comparable, additional comparative analyses are presented to evaluate the outcomes between VADER and FEEL-IT, to point out if similarities between positive and negative items could be outlined using the same textual communications. To this end, we considered matching positive VADER compound scores (> 0) with FEEL-IT-labelled *positive* sentiments, while non-positive VADER compound scores (≤ 0) with FEEL-IT-labelled *negative* sentiments.

2.3 Phase 3: extraction of lexical features through NLP methods

Lexical features (Saleem *et al.*, 2012) can help in the early detection of stress conditions of end-users, and thus, in the application context, of both organization workers (thus, end-users) and maintenance staff technicians. Among the different features, Lexical Diversity (LD) has been chosen based on (Zhou *et al.*, 2004). LD expresses the proportion of repeated words in a language sample and has been correlated with stress conditions (Alberdi *et al.*, 2016a; Vizer *et al.*, 2009), thus a reduction of cognitive performance. Different measurement methods of LD have been proposed (McCarthy and Jarvis, 2010; Woods *et al.*, 2023; Zenker and Kyle, 2021a). The first one is the Type-To-Token ratio (TTR), which is computed as the number of unique words with respect to the total words in a text. However, TTR is influenced by the length of the text when texts are too short.

Alternative methods and related indices (HDD, MATTR, MTLD, MTLD-MA, MTLD-BID, Root TTR, Log TTR, Maas, MSTTR) have been introduced to calculate LD overcoming the limits of a simple TTR calculation. Literature works comparing these approaches pointed out that, among these methods, Moving Average Type-to-token ratio (MATTR) seems to provide one of the "most stable" LD indices "maintaining a high degree of stability across all text lengths" (Zenker and Kyle, 2021b). In detail, MATTR calculates TTRs for a moving window of tokens from the first to the last token, computing a TTR for each window (Covington and McFall, 2010). MATTR is hence the mean of the TTRs of each window. MATTR is not affected by the text length, displaying a high degree of stability across a 50– 200 token range (Lei and Yang, 2020; Zenker and Kyle, 2021a).

The MATTR calculation has been performed for each end-user's maintenance request and for each description of the technical solution adopted by the technical staff, using textual communications in the dataframe pre-processed according to Section 2.1 methods. In particular, a Python script (rel. 3.11) has been written within the same Anaconda environment developed in phase 1 (Section 2.1), using Dataspell 2024.2.1.1 as the IDE, and implementing outcomes in the dataframe developed through "Pandas". In particular, Python library "textacy" (https://pypi.org/project/textacy/) has been used to comapute text statistics for textual communications, and toextend the functionality of the "spaCy" (https://pypi.org/project/spacy/) Natural Language Processing library (multilingual). "spaCy" has been used along with it_core_news_lg (Italian pipeline optimized for CPU, with components: tok2vec, morphologizer, tagger, parser, lemmatizer - trainable_lemmatizer, senter, ner). These libraries have been chosen due to their ability to manage texts in languages different than English, considering that the corpus of end-user's maintenance requests is written in the Italian language. Finally, the Python packages "Tool for the Automatic Analysis of Lexical Diversity" (TAALED) (https://lcr-ads-lab.github.io/TAALED/) and "lexical_diversity" (https://github.com/kristopherkyle/lexical_diversity) have been used to calculate LD indexes, which are reported in the supplementary material A.

After pre-processing actions, requests were processed by first assessing the number of total and unique words (which was also investigated by providing their histogram-based distribution), and then calculating MATTR for each of them. When the number of unique words and total words of a request is similar, MATTR tends to 1. Considering that MATTR is stable when the number of tokens is greater than 50 tokens (Zenker and Kyle, 2021a), we excluded the analysis sentences shorter than 50 tokens. The same pre-processing and processing actions could be also applied to calculate other LD indices. A low MATTR value indicates a reduced variety of vocabulary, suggesting a repetitive use of the same words. Since stress can affect an individual's cognitive and linguistic abilities, leading to less varied and more stereotyped communication. In high-stress situations, people may struggle to access a rich vocabulary, resulting in more repetitive language. This could explain why low MATTR values might be associated with high-stress conditions, but further specific research would be needed to confirm this correlation. Therefore, we assumed a plausible connection between low MATTR values and high-stress conditions, and thus comparisons and correlations between the indices obtained by these different occupant typologies have also been assessed.

2.4 Phase 4: definition of indexes to assess the workplace health status and comparison between the organization workers and technicians' perception

According to the paper aims, several indexes are proposed, aimed at identifying job-related stress associated with a specific *building complex*. This choice aims to assist facility managers in making strategic decisions about major workplace interventions, based on dynamic analysis of information on CMMS. Therefore, indexes are calculated using the outputs provided in Sections 2.1, 2.2 and 2.3, and considering each request (*i)* related to a type of occupants (*k)*, which can be end-users of the maintenance tasks (*EU*) or technicians of the maintenance staff (*TE*), or even subgroups within them, by filtering them by a specific *building complex* (*j*) of the building stock.

The main outputs derived from the application of methods reported in Section 2.1, Section 2.2 and Section 2.3 have been normalized by considering the occupancy load OL_j [persons/m^2], as shown in Table 1 (that is, the ratio between the number of workers of the building complex j and the related GFA_j). This choice allows to evaluate outputs by crowding conditions of the building complex, depending on its surface extent. In all the cases, as a result, when two *building complex*es report the same values for a given output, the higher the OL_j the lower the index. Results have been compared to stress possible insights differences between the normalization-based assessment methods.

Three different types of indexes have been considered, respectively relating to: (1) *workplace maintenance needs (WM)*, based on the analysis of the requests and on the contractor's response actions; (2) *perception of maintenance quality* (*PQ)* of the spaces, based on sentiment and emotion contents analysis, both related to end-users and technicians; (3) *job-stress detection* (*JB)* of both end-users and maintenance staff technicians, based on the analysis of the lexical feature.

WM indexes are based on the preliminary analysis of data described in Section 2.1, and thus they include:

- WM_M [requests / pp/m^2], which is the ratio between the number of maintenance requests (*MR*) for a specific *building complex*, and the related OL_j , as described by Eq.1. The higher WM_M, the higher the maintenance effort in terms of requests, but no insights on their complexity is assessed by this index.

$$WM_{R,k,j} = \frac{\sum_{i=1}^{n} MR_{k,i,j}}{OL_j} \tag{1}$$

- WM_T [days / pp/m^2], which is the ratio between a significant statistical distribution value of the time necessary to solve the requests (*TIME-TO-SOLVE_stat*) for a specific *building complex*, and the related OL_j, as described by Eq.2. The higher WM_T, the higher the effort to solve the request, but from an individual perspective since it relies on statistical descriptors of the *TIME-TO-SOLVE* data. To describe recurrent conditions, *TIME-TO-SOLVE_stat* can refer to the average (in case of normal data distribution) or median *TIME-TO-SOLVE* (in case of non-normal data distribution and to limit the impact of outliers). In addition, to consider critical conditions, the 95th percentile of *TIME-TO-SOLVE* distribution is also taken into account in this paper, thus excluding maximum times related to outliers.

$$WM_{T,k,j} = \frac{TIME-TO-SOLVE\_stat_{k,i,j}}{OL_j} \tag{2}$$

- WM_{Ttot} [days / pp/m^2], which is the ratio between the sum of *TIME-TO-SOLVE* of all requests for a specific *building complex*, and the related OL_j, as described by Eq.3. The higher WM_{Ttot}, the higher the total effort to solve all the requests from a global perspective, since it relies on the aggregation of *TIME-TO-SOLVE* data. Therefore, this value can be suitable to support the contractors in verifying which *building complexes* are associated with higher efforts for the technicians' teams working on them, and thus to provide insights on the staff deployment.

$$WM_{Ttot,k,j} = \frac{\sum_{i=1}^{n} TIME-TO-SOLVE_{k,i,j}}{OL_j} \tag{3}$$

In this paper, *WM* indexes are filtered by considering all the *EU* as well as the group *k*, who are the organization workers who submit maintenance requests to the contractor. According to their definition, *WM* indexes are ≥ 0.

PQ indexes are based on methods and outputs for sentiment analysis described in Section 2.3. They comprise two indices:

- $PQ_{Vtot,k}$ [1 / pp/m^2], calculated as in Eq.4, which refers to group *k* as end-users ($PQ_{Vtot,EU}$) and technicians ($PQ_{Vtot,TE}$). This index is based on the sum of VADER compound score (*VCS*) values, thus providing a complete overview of the sentiment-related conditions for the entire group of occupants.

$$PQ_{Vtot,k,j} = \frac{[(\sum_{i=1}^{n} VCS_{k,i,j})/(\sum_{i=1}^{n} MR_{k,i,j})]}{OL_j} \tag{4}$$

- $PQ_{V,k}$ [1 / pp/m^2], calculated as in Eq.5, which refers to group k as end-users ($PQ_{V,EU}$) and technicians ($PQ_{V,TE}$). $PQ_{V,k}$ is the ratio between a significant statistical distribution value of the VADER compound score-based assessment of requests (*VCS_stat*) for a specific *building complex*, and the related OL_j. *VCS* values range between -1 and +1. To describe recurrent conditions, *VCS_stat* can refer to the average (in case of normal data distribution) or median *VCS* (in case of non-normal data distribution and to limit the impact of outliers). Therefore, the 5th percentile of *VCS_stat* distribution is also taken into account in this paper, thus excluding possible outliers.

$$PQ_{V,k,j} = \frac{VCS\_stat_{k,j}}{OL_j} \tag{5}$$

According to *VCS* definition (compare also Section 2.3), the lower the *PQ*, the more negative end-users and technicians' perception of the claimed maintenance issue. On the other side, in this paper, the perceived emotions derived according to Section 2.3 (i.e. fear, anger, sadness, joy) are not considered to derive indexes, due to the qualitative character of this type of information.

JB indexes are based on methods and outputs for MATTR calculation described in Section 2.2. *JB* indexes include:

- $JB_{MATTRtot,k}$ [1 / pp/m^2], calculated as in Eq.4, which refers to group k as end-users ($JB_{MATTRtot,EU}$) and technicians ($JB_{MTATTRtot,TE}$). Considering the numerator of Eq.4, each MATTR value ($MATTR_{k,i,j}$) can have a maximum value of 1, as defined in Section 2.2. Therefore, the maximum numerator value will be the total number of maintenance requests.

$$JB_{MATTRtot,k,j} = \frac{\sum_{i=1}^{n}(MATTR_{k,i,j})}{OL_j} \cdot 1000 \tag{6}$$

- $JB_{MATTR,k}$ [1 / pp/m^2], calculated as in Eq.7, which similarly refers to group k as end-users ($JB_{MATTR,EU}$) and technicians ($JB_{MATTR,TE}$). $JB_{MATTR,k}$ is the ratio between a significant statistical distribution value of the MATTR values (*MATTR_stat*) for a specific *building complex*, and the related OL_j. *MATTR_stat* values range from 0 to 1. To describe recurrent conditions, *MATTR_stat* can refer to the average (in case of normal data distribution) or median (in case of non-normal data distribution and to limit the impact of outliers) MATTR value. Therefore, the 5th percentile of *MATTR_stat* distribution is also taken into account in this paper, thus excluding possible outliers.

$$JB_{MATTR,k,j} = \frac{MATTR\_stat_{k,j}}{OL_j} \tag{7}$$

According to their definition (compare also Section 2.2), *JB* indexes > 0, and the lower the *JB* indexes, the higher the possibility of job-related stress conditions.

In addition to *WM*, *PQ* and *JB* indexes, results from the Customer Satisfaction questionnaires on maintenance services were used to define the Maintenance Satisfaction Level from questionnaires of end-users $MSL_{tot,EU}$ [1 / pp/m^2], calculated as in Eq.8. The definition of this index follows the general rules of $JB_{MATTRtot,k}$, by adopting the sum of the end user satisfaction level expressed by the questionnaire vote ($msl_{EU,i,j}$ [-]) as Eq.8 numerator. The lower this value, the lower the level of satisfaction for maintenance tasks, and thus the higher the possible related stress conditions (Jaafar, 2021). In this work, it is worth remarking that this kind of data has been derived only for university workers and not for maintenance technicians, thus comparisons have been carried out by mainly considering $JB_{MATTRtot,EU}$ and $MSL_{tot,EU}$ for each building complex j.

$$MSL_{tot,EU,j} = \frac{\sum_{i=1}^{n} msl_{EU,i,j}}{OL_j} \tag{8}$$

All the indexes are calculated thanks to Python scripts in the Anaconda environment defined in Section 2.1.

3. RESULTS AND DISCUSSIONS

3.1 Preliminary analysis of maintenance requests dataset

Table 2 shows the number and relative percentage of maintenance requests and corresponding work orders during the monitoring period (2018-2023) per *building complex* (represented by ID).

Most requests (> 95%) relate to a limited set of *building complexes*, which are reported in bold, and are then selected for the next steps of the research. In particular, these requests are mainly linked with administrative offices and larger indoor educational & research buildings (compare intended use and GFA in Table 1). In this sense, more than 30% of requests essentially refer to the widest *building complex*, that is L&-047-006. These results are essentially due to the combination of the dimensional features and the attractive role of these *building complexes*, which typically host visitors and students who interact with components causing higher maintenance loads, as confirmed by the outcomes of previous works on similar intended uses (Gunay *et al.*, 2019; O'Brien *et al.*, 2017). On the contrary, minor and highly specialized *building complexes*, mainly serving outdoor areas (e.g. L6-047-012) or used as supplementary facilities (e.g. L6-047-017), are characterized by very limited percentages of requests.

Table 2: End-user maintenance requests by buildings complex according to the ID reported in Table 1. Complexes and data reported in bold are considered in the following analysis.

Building complex ID	**Number of requests**	**percentage of requests**	**Number of WOs**	**percentage of WOs**
L6-047-001	521	2.46%	508	2.50%
L6-047-002	**822**	**3.88%**	**804**	**3.96%**
L6-047-003	5	0.02%	3	0.01%
L6-047-004	**2605**	**12.29%**	**2500**	**12.31%**
L6-047-005	**5496**	**25.93%**	**5276**	**25.98%**
L6-047-006	**7220**	**34.06%**	**6983**	**34.38%**
L6-047-007	**2306**	**10.88%**	**2136**	**10.52%**
L6-047-008	**1974**	**9.31%**	**1868**	**9.20%**
L6-047-009, L6-047-010,L6-047-011,L6-047-013	42	0.20%	35	0.17%
L6-047-012	4	0.02%	4	0.02%
L6-047-014	23	0.11%	23	0.11%
L6-047-016	2	0.01%	2	0.01%
L6-047-017	177	0.84%	168	0.83%

In addition, Table 2 shows that about 96% of requests correspond to work orders, thus underlining that almost all requests correspond to an effective issue that had to be managed by the contractors. Only these requests are considered in the following steps, so as to have a complete overview of the process from both the university workers and the maintenance staff technicians, but this choice does not imply a significant reduction in the dimension of the analysed sample. Considering that during the monitoring period, the university workers were about 1000 people, this means that on average each worker experienced more than 20 maintenance issues.

Considering the work category, about half of the work orders refer to "electrical" equipment and systems (49%), while "heating and cooling" and "plumbing" systems refer to a similar percentage impact within the *work categories* (respectively, 28% and 23%). The relevance of "electrical" *work orders* is correlated to the functional obsolescence of the plants in view of recent and current needs for electrification improvements, especially in research areas.

Finally, for the whole building stock, the average time necessary to solve a *work order* is 2.28 days. The only "electrical" category requires a slightly higher average time for the intervention (2.58 days). The maximum recorded value is 80 days, necessary to solve electrical problems.

Figure 2 then shows the boxplot of the ratio between the time necessary to solve the request and the related expected contractual time, by *work category*. The compliance with the contractual time frame is represented by a ratio equal to 1 (indicated by the dashed vertical black line in Figure 2). This relationship provides evidence of the facility manager's ability to respond on time, potentially resulting in a lower level of stress related to maintenance activities, since the maintenance request is solved in the expected time. It can be noticed that all the *work categories* have a similar trend and close median values, and that almost all the corrective interventions have been made respecting the contractual time constraints (ratio ≤ 1). The 4% of work orders is out of the contractual time limit (ratio > 1), mainly representing outliers in the ratio distribution, as remarked by Figure 2. It is worth noting that only 0.4% of the whole work order number needed twice the expected time to be solved (ratio > 2), thus representing a negligible sample.

To explore which requests are affected by such conditions, an analysis of the distribution of the related *programmability* classes is performed along with an analysis of the textual description by end-users and of the short report by technicians. It is worth noting that most of requests (75%) refer to "*not delayable"* interventions, which have been generally supported by limited information by end-users, and thus needed more detailed in-situ inspection by technicians, e.g. for plumbing work order: "*I am Name Surname, and I am reporting a failure within Building 5, at level Q145 at the Y laboratory (comprised in the Z research area). We report the presence of abundant water on the floor coming from unknown origin. We require rapid intervention for safety reasons, to restart activities after the failure is solved. Thank you very much*". Nevertheless, a work order classified as "*not delayable"* is associated by very limited expected time for the intervention closure (e.g. up to a few or few hours), and thus the overall impact on business continuity has been considered as reasonable limited in the operational context of education facilities. The rest of outliers essentially refers to interventions associated with "long-terms" *programmability* and mainly needing specific components and equipment, thus being affected by market availability and procurement plan.

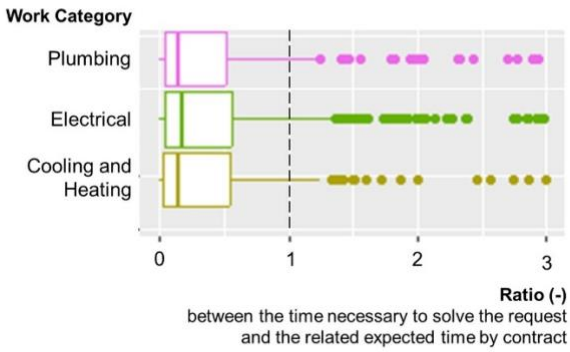

Figure 2: Boxplots of the ratio between the time to solve the request and the related expected contractual time, by work category. The box represents the interquartile range, that is, the range between the 25th percentile (lower bound of the box) and the 75th percentile (upper bound of the box). The line inside the box indicates the 50th percentile (median value). The compliance with the contractual time frame is represented by a ratio equal to 1 (indicated by the dashed vertical black line).

3.2 Users' and technicians' perception based on the assigned priority/programmability classes and on sentiment and emotion analysis

In this section, the analysis refers to those end-users' requests that were translated into *work orders*, since they comprise the *programmability* labels and the final textual description of the intervention by technicians' written communications.

Considering data on the whole building stock, end-users associated 17.4% of their requests with the "*emergency*" label (representing the maximum *priority*). 0.8% of requests are labelled as "*urgent*". Therefore, most of the requests (81.8%) are labelled as "*not urgent*" (representing the lowest *priority*). In general terms, end-users' requests widely comprise expressions such as "(very) urgent/urgently", "as soon/quick as possible", "immediate action required", as well as they point out that the failure involves more than an area/department or a wide group of workers (compare with Section 3.2. examples). On the contrary, 7.7% of *Work orders* (hence associated with these requests) were labelled by technicians as "*Not delayable*" actions, while 41.2% of them were associated with the "*short-term*" *programmability* label. These *programmability* labels essentially correspond to end-users' "*emergency*" *priority* labels, according to the assumptions expressed in Section 2.1. 36.7% of work orders were labelled as "*medium-term*" (which corresponds to "*urgent*" *priority*), while 14.4% as "*long-term*" *programmability* actions (corresponding to "*non urgent*" in the *priority* scale).

The analysis of the *programmability* attributions underlines a scarce consistency with the *priority* attribute provided by the users, especially for the highest *priority*/*programmability* classes. First, the perception of the end-users about the gravity of certain issues can be influenced by their limited technical knowledge. This leads to a general undervaluation of maintenance issues, as shown by the low number of requests characterized by "*emergency*" *priority* with respect to work orders classified with "*not delayable*" and "*short-term*" *programmability* labels. On the other hand, the very high number of end-users' requests labelled with "*not urgent*" *priority* could be also probably due to the limited number of levels adopted by the facility manager in the *priority* scale that is proposed to the end-users. Nevertheless, it is necessary to consider that the technical staff, in particular, situations (e.g. especially when the number of contemporary requested actions is very relevant and significantly more than usual), can be not enough to manage all the requests. Then, the attribution of the *programmability* of the action could be also influenced by the dimension of the staff involved.

Figure 3 shows the sentiment polarity distribution by *priority* and *programmability*, i.e. the distribution of the computed VADER compound score (*VCS*) for end-users and technicians, vertically aligned according to the classification similarities defined in Section 2.1 (see vertical dashed lines).

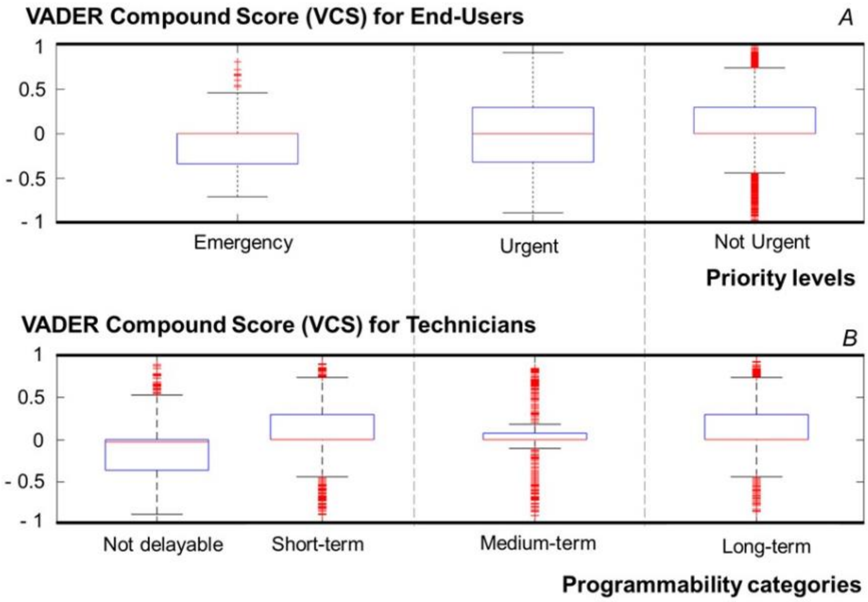

Figure 3: Boxplot distribution of the VADER compound score (VCS) for (A) end-users and (B) maintenance technicians, grouped respectively by priority levels (Emergency, Urgent, Not Urgent) and programmability categories (Not delayable, Short-term, Medium-term, Long-term), as defined in Table 1. The box represents the interquartile range (IQR), spanning from the 25th to the 75th percentile; the horizontal line within the box indicates the median (50th percentile). Whiskers extend to 1.5 times the IQR, and red crosses denote outliers. Vertical dashed lines highlight the conceptual correspondence between priority and programmability classes, as discussed in Section 2.1.

As expected, the lowest negative *VCS* were found for interventions labelled by the users as "*emergency*" (Figure 3-A). In this case, almost all end-users' maintenance requests are characterized by values included between -0.5 and 0. Requests labelled as "*urgent*" *priority* are characterized by *VCS* between -0.5 and 0.5, and, in particular, half of the values are positive. This output suggests that the 3-level scale used by the facility manager to support end-users' *priority* attribution could not completely express the real perception of the end-users. Moreover, as expected, requests labelled as "*not urgent*" are characterized by positive values, essentially comprised between 0 and 0.2. Results could be affected by the sample dimension by priority class, indeed, and thus the number of outliers in the "*not urgent*" class could appear as more relevant than in the other ones. Nevertheless, outliers in this category (*VCS* > 0.5, thus more positive ones) could be also explained by the fact that some requests include acknowledgements to stimulate a fast reply by the technicians (e.g. politeness and courtesy expression such as "please kindly…", "Please notice that…").

The *programmability*-related *VCS* distribution (Figure 3-B) seems to follow the same rationale as the *priority*-related one. As expected, "*not delayable*" *programmability* classes are characterized by the same lower *VCS* values noticed for the "*emergency*" *priority* class. The "*short-term*" *programmability* class generally seems to provide a shift of VCS distribution towards positive values with respect to the "*not delayable*" class. Nevertheless, for "*medium-term*" and "*long-term*" *programmability* classes, VCS distributions are quite similar to those of "*short-term*", essentially underlining that median technicians' *VCS* is still close to neutral (that is balanced) technicians' sentiment.

Finally, the performed correlation analysis based on the Kendall method shows that no correlation exists between the time necessary to solve the issue and the *VCS*. The tau-kendall score is 0.058 (p-value < 2.2e-16).

Figure 4 shows the *VCS* distributions via boxplot representation, related to end-users' and technicians' texts, by building complexes. In general terms, most *VCS* median values are close to 0, especially for the technicians, thus underlining neutral sentiment. *VCS* < 0 are generally limited to the 25th percentile of data, while almost half of written texts seem to be related to neutral to positive sentiment. Limited differences can be highlighted in Figure 4 among the different buildings' complexes. Different reasons could be considered under such a result. First, the different sample dimensions in terms of maintenance requests could hence impact the final distribution of *VCS* values, and then the presence of outliers, too. Nevertheless, matching Figure 6 and Table 2, it can be roughly noticed that building L6-047-002, which is affected by the lowest number of maintenance requests, is also characterized by more "positive" quartiles of the end-users' *VCS* and a higher related median *VCS* in respect to the other buildings complexes, thus denoting a sort of lower perceived severity of maintenance issues, as expected. From the technicians' perspective, *VCS* values are more condensed towards neutral values for building L6-047-006, which owns the highest *GFA* and number of work orders.

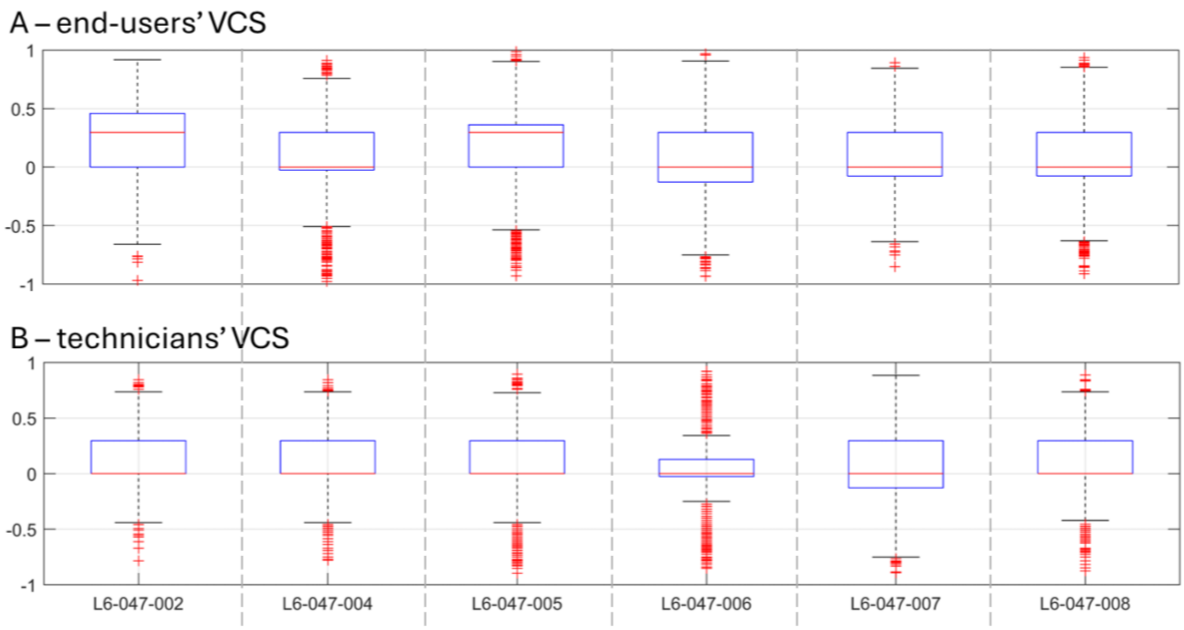

Figure 4: VADER compound score (VCS) boxplot distribution for end-users (A) and maintenance technicians (B), by building complex ID (Table 1), highlighted by vertical dashed lines. Crosses represent outliers.

FEEL-IT has been used to process the texts and extract the basic emotions of both end-users and technicians, as represented in Figure 5.

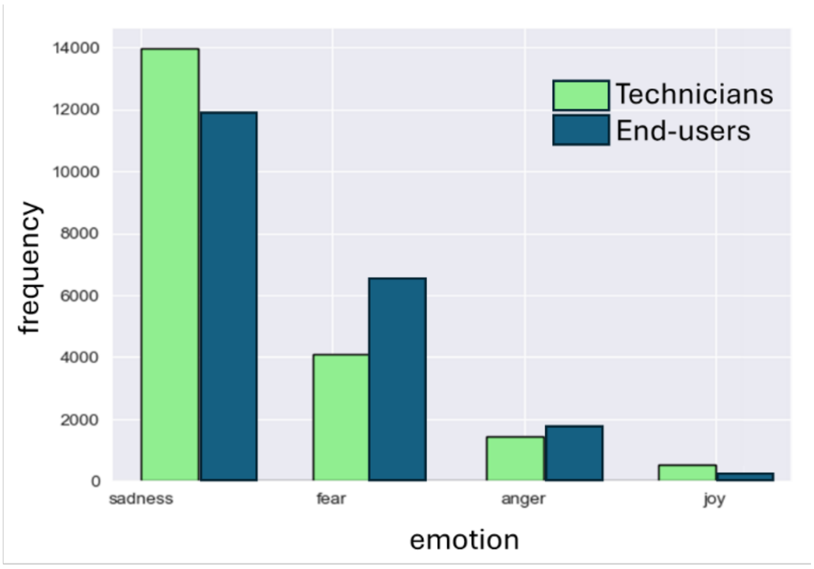

Figure 5: Frequency of sentences labelled with one of the four Ekman's basic emotions (fear, sadness, anger, joy) for end-users' and technicians' groups.

Considering that maintenance requests are written to give evidence of a specific issue, as expected, texts labelled with the emotion "joy" are residual, and probably due to sentences including acknowledgements and politeness and courtesy expressions to stimulate the attention of the technical staff, as discussed above. A representative example of this kind of request, conveying a sense of friendliness and positivity, follows: "*Good morning, I am writing to you from department X, building 3A, level 2 (where the office of Prof. Y is located). In rooms 004, 005, 006, I noticed some neon lights that maybe need to be replaced, so we would kindly require the intervention of a technician from the maintenance service, when possible. Thank you! Have a nice day and good work, Name, full address and phone number*".

For the end-users, more than 6000 requests were labelled with the "fear" emotion and about 2000 with the "anger" emotion. It can be assumed that these requests can express the stress condition caused by the issues with the equipment, involving, e.g. the impossibility of concluding planned actions or suffering from additional risks from the failure. Two relevant examples follow, focusing on end-users' requests:

- for "fear"-classified request, conveying a sense of urgency and disruption along with safety concerns: "*Building complex L6-047-006, building 5, level Q145. During last week, bathrooms of the laboratory of department X were overflowed, and this event should have caused a problem with the electrical system. We are suffering blackouts in bathrooms since yesterday morning, while, since this morning, blackouts also involve our office rooms. Moreover, on Friday, an unjustified fire alarm was triggered in the same area. Last night, inexplicably, the emergency lights of the old system came on despite the fact that the electricity had not gone out. Name Surname, mobile phone number".*
- for "anger"-classified request, expressing frustration and urgency due to the unresolved issues for daily working life in the workplace: "*Building complex L6-047-006, building 3B, level Q145, department X. Unfortunately the intervention of the technician you sent did not solve the problems of minimum heating in the department X offices: we are literally freezing to death! I mean, not only teachers, but also the many students who attend the area! This morning the students, although covered by their coats, complained a lot. I therefore ask you to intervene as a matter of urgency. Name Surname, mobile phone number.*"

"Sadness" is the prevalent emotional condition detected, indeed about half of the maintenance requests are associated with this emotion. It comprises different emotional states from mild disappointment to extreme despair and anguish (i.e. disappointment, discouragement, distraughtness, resignation, helpness, hopelessness, despair, grief, and sorrow). A relevant example for the end-users follows, conveying a sense of worry and helplessness along with urgency and concerns due to the deteriorating conditions of the plants, also in view of possible obsolescence issues: "*Hello, I am writing to report a problem in room X a the department Y greenhouse, where we have our plants. Given the current temperatures, we had to turn on the cooling system in the laboratories. We wanted to clean the filters of the fan coils, but they practically no longer exist: they seem to "crumble". I therefore ask to be able to replace them as soon as possible. Thank you!*"

Results related to technicians' emotions show a similar trend to those of end-users, with a slightly higher frequency for "sadness" emotion. In this case, it is relevant to note that fear percentage increases in technicians' texts. A relevant example from the technicians' communication follows, conveying a sense of "fear" related to the necessity to always ensure safety, and connected with possible frustration due to multiple schedules of interventions, associated with disruption of services: "*From the checks carried out, it is found that there is a fault in the fan coil motor which causes the detachment of the switch in the electrical panel located on the attic floor. Decisive interventions were firstly scheduled and completed. Then, the final intervention has been carried out, comprising the replacement of the complete engine and restoration of the functionality of the fan coil in the area. Please note that, for safety reasons, it is absolutely forbidden for all unauthorized personnel to carry out operations on live systems*."

Finally, an additional comparison between VADER and FEEL-IT outcomes considering positive and negative items has been reported in supplementary material B. Results indicate a slight prevalence of positive correspondence (55%), i.e. the same direction for both indicators, which could be essentially due to the high number of sentences labelled with the sadness emotion (which is not always interpreted as a negative emotion) by FEEL-IT.

3.3 Lexical diversity analysis of the end-users' maintenance request and the technical staff solutions

End-users' maintenance requests in the whole building stock, analysed by MATTR methods as defined in Section 2.2, are characterized by a mean length of 67.8 words, ranging up to 181 words (1st quartile: 54; 2nd quartile: 62, 3rd quartile: 75). The number of unique words ranges from 20 to 117, with a median value of 48 words and a mean value of 50.9 words. The technicians' texts have a similar trend, being characterized by a mean length of 66.1 words, ranging up to 177 words (1st quartile: 54; 2nd quartile: 61, 3rd quartile: 70). Unique words in technicians' texts range from 22 to 89, with a median value of 44.7 words and a mean value of 44 words. The number of total and unique words in assessed texts has a non-normal distribution. To provide a more detailed overview on the matter, Figure 6 resumes these values in terms of cumulative function distribution (cfd), for data on end-users (Figure 6-A and B) and technicians' (Figure 6-C and D) texts. Figure 6 hence offers an overview on the length and thus verbosity of the communication, by the number of words in requests (Figure 6-A) and technicians' reports (Figure 6-C), along with the complexity of the claimed matter and richness in vocabulary, bu number of unique words from end-users (Figure 6-B) and technicians' (Figure 6-D) texts. In general terms, similar trends between the occupant-related samples, although it could be claimed that technicians' texts are characterized by a slightly lower number of unique words. In this sense, Figure 6 reflects the structure of text contents, as well as the specific information reported within each sentence depending on the two types of "writers". Some relevant examples are discussed in the following.

End-users' requests are generally structured by: introduction and initial greetings (e.g. "*Dear maintenance service staff, sorry to bother you, but we wanted to let you know that…*"); data on the localization of the maintenance request, although related building complex ID has been generally not provided by the end-users ("…*here, in our department X area on the -1 floor of the central building (Building 1, Wing C, Basement floor)…*"); description of the maintenance request ("…*we regularly experience daily power outages. ...*"), by often adding other details ("…*Specifically, this seems to happen either during the night or early in the morning. Outside, in the main electrical panel (located in front of the bar corridor, to be precise), we find the circuit breaker lowered upon our arrival, corresponding to the electrical system marked by the label "LEFT SIDE." Similarly, last week a circuit breaker tripped in the electrical panel in front of the department y area. ...*"); final thanks and greeting ("…*Thanks indeed for the precious support and hope to hear from you soon, Name Surname*"). Longest requests (as the one in the previous example) were widely characterized by redundant greetings and/or a description of minor details on the failure, remarking that the failure is widely affecting other workers or divisions.

On the contrary, technicians' texts are generally structured by describing the interventions, in terms of activities and localization, and by often repeating the same technical words. As a relevant example, the end-users' request (original number of words in Italian: 92, with 68 unique words): "*Low temperature has been reported, since last week, in rooms relating to X area on the ground floor of Faculty Y, building 1. This morning, it was found that many heaters do not reach the normal operating temperature. Some of them, placed in bathrooms and corridors at first floor of the same building, were completely cold. Misfunctioning was also reported on elements at first and second floors of Faculty Y Building 2. We therefore ask for an inspection of the heating systems of the whole building Complex. Thank you.* " corresponded to the technicians' text (original number of words in Italian: 62, with 42 unique words): "*Building 2 – After inspection, the heating system serving the building is functional and comfortable temperatures are found in the rooms; Building 1 - After inspection, the heating system serving the building is active, work has been carried out to repair and restore the functionality of the pump serving the overall circuit, with final control of the heating bodies in the various rooms*".

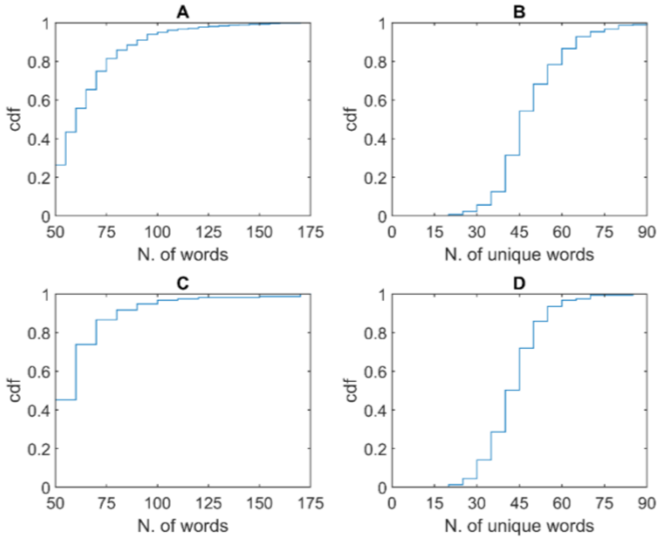

Figure 6: Cumulative density function for the number of words and unique words in end-users (respectively A and B) and technicians' (respectively, C and D) texts.

In the following, the comparison between MATTR values of end-users and technicians have been performed considering only those requests that correspond to work orders, since they include final textual communications written by the technicians to describe the performed maintenance intervention. Figure 7 explores if stress levels detected by textual communications related to the same work order seem to be correlated between end-users and technicians. Therefore, Figure 7 graphically traces the correlation between MATTR values for technicians (TECH, Figure 7-A) and End-users (EU, Figure 7-B) along with the scatter plot (Figure 7-C) of them, considering points are single maintenance requests. The Pearson's Correlation Coefficient is also shown in Figure 7-C.

From a general point of view, MATTR values are non-normally distributed, pointing out the highest frequencies for the highest MATTR values for both the typologies of occupants. MATTR distribution for technicians is characterized by slightly lower values (1st quartile: 0.78; 2nd quartile: 0.86; 3rd quartile: 0.92) with respect to that of end-users (1st quartile: 0.83; 2nd quartile: 0.89; 3rd quartile: 0.93), as graphically remarked by Figure 7-C pairs. Moreover, Figure 7 suggests that technicians' MATTR-based stress is slightly higher than the one of end-users, since MATTR is generally slightly lower. It could be argued that maintenance requests are performed by a high number of occupants, while maintenance activities are performed by a lower number of technicians. Nevertheless, each textual communication from technicians on the performed actions to solve the related maintenance issue is quite short and based on technical details, according to the examples given above. Therefore, the possibility to repeat words is limited, while having an extensive (technical and action-related) vocabulary with still a limited text length. Finally, Figure 7-C also shows that the Pearson's Correlation Coefficient does not seem to point out a strong correlation between MATTR values for the two categories of occupants (p-value < 0.01). Nevertheless, this result suggests that requests characterized by higher potential stress for end-users could not have the same "stress" effect on technicians. Reasons under this result could be essentially related to the possible differences in perception of maintenance activities by these two categories of occupants.

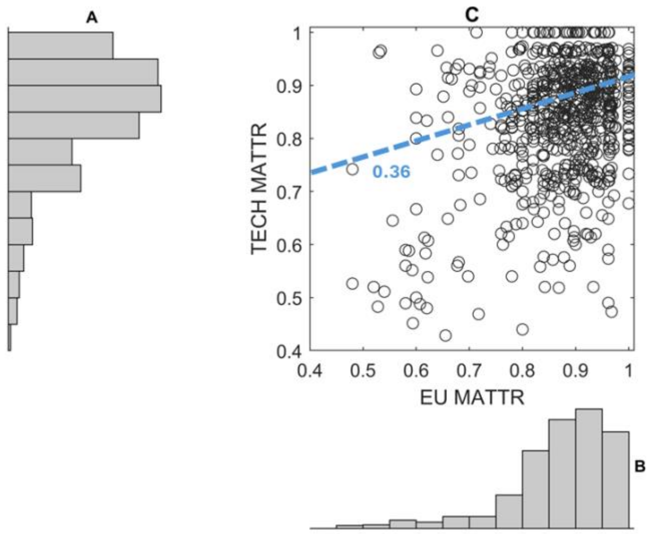

Figure 7: MATTR values distributions for: A) technicians (TECH); B) End-users (EU); and C) scatterplot of pairs for the maintenance requests, including Pearson's Correlation Coefficient.

3.4 Evaluation of the indexes on workplaces health status

Figure 8 reports the evaluation of the indexes proposed in Section 2.4, thus tracing, per each building complex, the overview of: (1) the workplace maintenance needs (*WM* indexes); (2) the end-users and technicians' stress and emotion states (*JB* and *PQ* indexes) derived from textual communications analysis; (3) the end-users' satisfaction on maintenance tasks (*MSL* index) derived from QASO questionnaires, and only related to university workers. While providing normalization by the different OL_j, some indexes have been multiplied or divided by 1000, to ensure calibration of visualization with respect to the numerator of each equation. The following analysis of results mainly focuses on the worst conditions highlighted by the indices (red coloured cells in Figure 8), since they relate to the critical scenarios with priority of interventions for maintenance management optimization.

From a general perspective, considering Table 1 data, Figure 8 points out two main relevant scenarios for intervention management and priority optimization to be considered.

The first one concerns the largest building complexes, characterized by lower occupant loads, i.e. L6-047-006 and L6-047-004, which relate to educational and research intended uses, and widely host large single rooms and corridors (especially for L6-047-006). They denote more critical conditions in maintenance request number and intervention timings, as pointed out by *WM*-related indexes, while users' perception issues, expressed by *JB* and *PV*-related indexes, are less relevant. The complexity due to the large surface of the building complexes implies a greater number of requests, due to the relevant extension of building equipment to which failures are related. At the same time, workers seem to be less "worried" and "stressed" about the failures. This kind of result seems to be also confirmed by the higher *MSL* values in respect to the other building complexes, which remarks a general higher satisfaction of workers on the way maintenance activities are carried out. Since building complexes are very large, workers could be spread out across different areas. Due to varying occupancy schedules and the extensive size of the buildings, the overall schedule is not intensive or continuous. This reduces the simultaneous, prolonged massive presence of many workers in the same space, and they seem to be less likely to be affected by the same maintenance issues. Similarly, it could be assumed that maintenance technicians could expect high engagement and workload in these scenarios, thus pointing out less critical (i.e. highest *JB* and *PV* values) in the sample.

On the contrary, the second group of scenarios is represented by the smallest building complexes, characterised by the highest occupant loads, e.g. L6-047-002 and L6-047-007. In this case, the quantity of maintenance requests has the less critical impact within the building stock, being easy to manage over time (lower WM_T values). Nevertheless, the impact on occupants' perception, expressed by total values of *JB* and *PV*, seems to be more critical than in the other buildings complexes, and this could be due essentially to the fact that occupants daily face overall building plant failures in the same spaces, within a more limited area. *MSL* values for these buildings are the lowest ones for the whole building stock, thus remarking how the lower aggregated level of satisfaction for maintenance tasks of workers seems to confirm *JB* and *PV* results.

Moreover, *MSL* results seem to be generally consistent with the textual communications analysis using MATTR and VADER, respectively used to calculate *JB* and *PV* values, for workers. Nevertheless, the case study application remarks critical opposite conditions, while slighter differences could be seen for intermediate scenarios which do not belong to the two main relevant scenarios defined above.

Some additional insights could be linked to *PV* outcomes, indeed. As pointed out by median values $PQ_{V,EU}$ and $PQ_{V,TE}$, which are generally close to 0 for all the buildings complexes, both end-users and technicians' sentiments tend to be quite neutral/balanced but slightly positive, confirming general results provided by Section 3.3 analysis. A possible analysis of these results could relate to a sort of perceived utility of the claimed requests by end-users and of their maintenance work by technicians. Nevertheless, the 5th percentiles of $PQ_{V,EU}$ still confirm a quite negative sentiment which is greater considering the end-users. In this sense, results confirm the reliability of the proposed indexes to describe users' perception issues.

In view of the above, the proposed indexes provide some suggestions for the optimisation of the operational tasks on widespread building equipment, depending on the different issues affecting workplace health status. Maintenance needs in larger buildings should be continuously monitored in a priority way, to check that *WM* indexes do not exceed current values. Meanwhile, to move towards the reduction of workers' stress, priority actions could be performed towards administration offices (i.e. L6-047-002), trying to balance requests and make end-users more aware of maintenance efforts and performed interventions.

ID	WM_M requests / pp/m² / 1000	WM_{Ttot} days / pp/m² / 1000	WM_T Median (95th prc) days / pp/m²	$PQ_{Vtot,EU}$ 1 / pp/m²	$PQ_{V,EU}$ Median (5th prc) 1 / pp/m²	$PQ_{Vtot,TE}$ 1 / pp/m²	$PQ_{V,TE}$ Median (5th prc) 1 / pp/m²	$JB_{MATTR,EU}$ Median (5th prc) 1 / pp/m²	$JB_{MATTRtot,EU}$ 1 / pp/m² / 1000	$JB_{MATTR,TE}$ Median (5th prc) 1 / pp/m²	$JB_{MATTRtot,TE}$ 1 / pp/m² / 1000	$MSL_{tot,EU}$ 1 / pp/m² / 1000
L6-047-002	23.4	27.8	3.1 (186.5)	7.2	8.4 (-10.9)	0.8	0.0 (-10.0)	26.8 (23.3)	20.8	28.5 (23.8)	21.4	19.2
L6-047-004	1157.9	1514.0	68.71 (3137.6)	21.3	0 (-218.8)	11.7	0.0 (-170.2)	428.5 (318.8)	1009.6	445.7 (362.1)	1037.5	286.0
L6-047-005	1234.2	1243.6	26.7 (1337.7)	39.8	66.5 (-85.8)	6.8	0.0 (-85.8)	212.9 (179.8)	1070.8	224.7 (190.1)	1107.0	253.7
L6-047-006	3299.6	3283.3	47.3 (2678.6)	8.8	0.0 (-209.9)	16.3	0.0 (-155.6)	457.6 (381.3)	2975.5	457.6 (381.3)	2991.8	693.5
L6-047-007	207.1	158.6	7.065 (376.0)	7.4	0.0 (-42.7)	2.4	0.0 (-39.6)	86.4 (75.5)	178.9	90.0 (75.8)	181.3	51.0
L6-047-008	369.0	427.7	13.3 (1354.8)	15.7	0.0 (-92.3)	8.0	0.0 (-88.5)	177.8 (146.0)	313.6	187.1 (159.5)	324.1	71.2

Figure 8: Indexes evaluation by buildings complexes, providing normalization by occupancy load OLj. Colour scales provide general insights by column qualitatively ranging from best (green) to worst (red) conditions expressed by each index.

4. CONCLUSIONS AND FINAL REMARKS

Maintenance communications from building users are valuable sources for data-driven analysis of building health status. In fact, they can provide data on the effectiveness of maintenance activities, and on the end-users' perception of maintenance tasks, including possible insights on how maintenance needs can affect workplace quality in terms of satisfaction, well-being and possible stress levels for both end-users and technicians.

This paper takes advantage of textual data stored in CMMSs to provide analysis for facility managers and organizations aimed at enabling better-informed decisions on maintenance priorities based on both objective data and workers' feedback. Therefore, the methodology and the indicators proposed by this paper could help decision-makers of large organizations in the definition of major intervention to improve the quality of the workplaces, and consequently workers' productivity. In particular, considering the research questions of this contribution, the following key remarks arise.

- *(RQ1): Is it possible to identify critical conditions in the building stocks thanks to the combination of the number and timing of the request and of the users' perceptions?* The proposed indexes can provide useful insights to have a general overview of the workplace health status in terms of maintenance needs (how many corrective actions are needed and how long the process is) and related users' perception (which is noticed stress level in requests, and which is the quality perception). The current work aggregates data by building complex, since this is the main level managed by the used CMMS in the case study and because this paper focuses on the analysis of maintenance requests concerning widespread building plants that are implemented and managed at this scale. The normalization of indexes by occupancy load per building complex is useful to trace differences among workplaces within the organization. Nevertheless, the proposed methodology can be replicated by considering the composing areas within each building complex, while other normalization procedures could be tested in other case studies. This kind of application would be oriented towards a more detailed overview of critical/priority areas where to intervene, as well as would be prone to support assessment tasks considering other categories of maintenance requests correlated with local equipment and building components.

- *(RQ2): Is it possible to detect the perception of the workplace quality using sentiment and emotion analysis, and to provide insights on the impact of workplace health status and related possible stress conditions of workers using lexical diversity indexes, based on information stored in CMMS?* For the first time at the authors' knowledge, this paper provides specific indexes on quality perception (based on sentiment analysis) related to maintenance tasks in workplaces and to the related possible levels of job-related stress (based on MATTR analysis, which can provide a first, rough estimate of stress issues, especially for the lower index values). In this sense, comparison between these indexes with results from workers' satisfaction questionnaires on maintenance tasks have been also carried out to preliminarily verify the consistency of proposed indexes with insights from Quality Assurance surveys (which are consolidated tools for building managers). Higher stress levels detected by lexical features and lower quality perception outcomes calculated by sentiment analysis are generally confirmed by lower workers' satisfaction levels coming from questionnaire data. This outcome encourages future applications of the proposed methodologies. Nevertheless, results underline that, although these indexes can provide a general overview of the matter by group of occupants (i.e. end-users versus technicians), additional analysis should be carried out to have a more complete and detailed evaluation of causes for workplace quality, such as those relating to the history of the buildings, of its intended uses, of the occupancy level over time, as well as to other sociological and psychological topics that were not taken into account in this paper since they cannot be managed via CMMS.
- *(RQ3): Is there any difference between the perceptions of end-users (i.e. organization workers) and of technicians on the maintenance issues, especially considering timing and priority, given their different levels of knowledge of maintenance tasks?* Perceptions of end-users and maintenance technicians seem to be different. On one side, differences can be noticed according to the assigned *priority* and *programmability* of interventions. In this sense, end-users seem to be not aware of the meaning of some *priority* classes, due to their limited level of knowledge of the matter. Thus, *programmability* assigned by maintenance technicians is different from the end-users' *priority*. On the other side, the question has been also replied to by introducing VCS-based indexes. In general terms, the technicians' VCS-based perception seems to be more neutral than the one of the end-users, as expected. In fact, technicians are more aware of the specific technical context in which interventions are carried out, and thus this can limit the frequency of more negative sentiment-related conditions. Nevertheless, possible impacts on possible stress levels seem to be similar for these two categories, as also confirmed by the comparison between *job-stress detection* indexes exploiting lexical diversity indexes. In view of these results, it could be claimed that a potential loss in alignment of expectations and perception of interventions quality between these two workers' categories could exist. In this sense, additional analysis can be carried out by identifying specific homogeneous subgroups of occupants in each analysed space, as well as short interviews towards end-users and technicians could be performed and related feedback (e.g. using structured online forms) could be collected to better provide refined validation of automatic quantitative assessment results reported in this paper methods. Nevertheless, data anonymization approaches should be used indeed, as well as the assessed sample could be not small enough to allow the identification of the specific workers. The need to balance privacy and control issues with respect to the stress of workers should be accurately investigated, taking into account GDPR requirements and privacy restrictions (EU 2016/697), as also suggested by consolidated works (O'Neill and Carayon, 1993). Furthermore, building owners, having their own maintenance quality expectation and perceptions, should be also involved in

the loop, since they are the third party involved in tasks management and mainly take advantage of low stress levels of workers (Zubair *et al.*, 2024).

Despite the results obtained, which underline the capabilities of the proposed methodology and indexes, and the large dimension of the dataset, this paper has some limitations.

First, the dataset has been extracted from a specific CMMS, which relates to a specific organization (i.e. a university in Italy) and a specific building stock, realized in certain periods and having certain technical and geometrical characteristics. Therefore, applications in different contexts, in terms of, e.g. organization types, workers' typologies, geographical areas, linguistic variations, types and dimensions of building stocks, are strongly encouraged, also in view of positive assessment results with respect to the given research questions. In this sense, future applications could also comprise buildings where requests could not come only from workers, but also from non-workers as significant end-users (e.g. visitors, customs), using the same techniques and thus exploring quality perception differences among them. Some relevant examples could be related, but not limited to schools, indeed (by comprising students), wide commercial buildings, train and metro stations, and airports. From this point of view, the approach could also take advantage of additional surveys, short interview and feedback from end-users in correlation with maintenance needs and related quality perception, also by group wide enough to ensure relevance of insights and manage data properly. In fact, this kind of actions would imply the definition of structured forms, ensuring anonymous treatment of data, according to GDPR (2016/679 EU Regulation). Moreover, this paper considers that the facility management process on maintenance requests ends with the technicians' textual communication once they finished the maintenance activity. Post-intervention feedback from workers could be relevant to be included in the assessment look, to explore, for instance, if and how perception changes from the initial request. Similarly, an evolutionary evaluation process (calculating the indexes over different time spans) could be useful to evaluate if the overall maintenance management systems have been improved thanks to the insights from the assessment process.

Second, maintenance requests are short texts, then the values of some of the indicators can be affected by the dimension (length) of these texts. In particular, the measured lexical diversity, even if corrected with the procedure suggested by previous approaches (McCarthy and Jarvis, 2010), could be high if compared with the value referring to longer texts. Despite consolidated works on its use in maintenance requests analysis exist, it should be noticed that VADER compound scores are influenced by the characteristics of the texts, too. Maintenance requests are typically characterized by negative or nearly 0 compound scores, in view of the impact of failures on the proper functionality of buildings and organizations. Finally, it is necessary to consider that the texts written by the technicians are characterized by a limited vocabulary and without expressive words, with a clear influence on the calculation of these scores. Their support for decision-making could be hence less relevant than those given by textual requests written by workers. Additional effort should be hence devoted at improving the reliability of lexicons used by sentiment and emotion analysis (including the ones based on VADER used in this paper) moving from general purpose ones to adaptations relying on specificities of maintenance activities in buildings (and, thus, the related *Thesauri*), also depending on the main features of the building stocks. Similarly, this paper assumes a plausible connection between low MATTR values and high-stress conditions which are consistent with literature works and seems to be confirmed by the comparison with workers' satisfaction questionnaire results. In this sense, this pilot verification suggests that further efforts towards this direction could be promoted in future research. Nevertheless, the authors are aware that the link between low MATTR and high stress could remain still speculative. Therefore, additional studies would be required to directly empirically confirm this hypothesis and relationship, also including validated tools (e.g. stress self-assessments, physiological data, or other validated stress questionnaires) and by extending comparisons also the maintenance staff technicians. the

Despite these limitations, this paper provides the basis for the definition of automatic procedures to detect the point of view of users about the health status of buildings, dynamically processing information collected during their life cycles and including issues on the workplaces' quality perception and effects on the workers' activities. The findings are indeed particularly useful for facility managers and organizations to prioritize maintenance interventions based on both objective data and workers' perceptions.

The integration of qualitative and quantitative data in workplace health assessments offers new research opportunities for improving facility management strategies and enhancing worker well-being. In this sense, future works can use the proposed data-driven framework while moving towards an improvement of workplaces, thanks to more effective maintenance management. In particular, the proposed indexes could support the identification of priority spaces within the whole organization where to primarily intervene, and they could be supported by additional assessment processes based on workers' personal beliefs on workplaces quality, from the sociological, physiological, and psychological standpoint, too. In addition, future efforts could be oriented towards including results from targeted interviews and feedback in correlation with maintenance needs and related quality perception. Indeed universities, as well as other public administrations, typically administer surveys related to the efficiency and effectiveness of services, to users' satisfaction and to work-related stress (e.g., for the Italian context as relevant for the case study applications, see for instance (Mucci *et al.*, 2015)). These surveys help collecting specific information, such as the perceived well-being of staff, including assessments of the work environment and working conditions, and their level of satisfaction regarding various university services. The same general rationale concerning privacy regulation must be considered in this kind of activity.

ACKNOWLEDGEMENT

This study has received funding from the Project "Occupant-based DIGITal predictive MANagement to improve the built environment" DIGITMAN – Project code 2022JMRX2A – CUP I53D23001770006, funded under the National Recovery and Resilience Plan (NRRP), Mission 4 - Component C2 - Investment 1.1, "Fondo per il Programma Nazionale di Ricerca e Progetti di Rilevante Interesse Nazionale (PRIN)" - Call for tender no. 104 of 2/2/2022 and Concession Decree no. 961 of 30/06/2023 of the Italian Ministry of University funded by the European Union – NextGenerationEU.

REFERENCES

Ahuja, R. and Banga, A. (2019), "Mental stress detection in university students using machine learning algorithms", *Procedia Computer Science*, Elsevier B.V., Vol. 152 No. October, pp. 349–353, doi: 10.1016/j.procs.2019.05.007.

Alberdi, A., Aztiria, A. and Basarab, A. (2016a), "Towards an automatic early stress recognition system for office environments based on multimodal measurements: A review", *Journal of Biomedical Informatics*, Elsevier Inc., Vol. 59, pp. 49–75, doi: 10.1016/j.jbi.2015.11.007.

Alberdi, A., Aztiria, A. and Basarab, A. (2016b), "Towards an automatic early stress recognition system for office environments based on multimodal measurements: A review", *Journal of Biomedical Informatics*, Elsevier Inc., Vol. 59, pp. 49–75, doi: 10.1016/j.jbi.2015.11.007.

Alruily, M. (2023), "Sentiment analysis for predicting stress among workers and classification utilizing CNN: Unveiling the mechanism", *Alexandria Engineering Journal*, Faculty of Engineering, Alexandria University, Vol. 81 No. August, pp. 360–370, doi: 10.1016/j.aej.2023.09.040.

Assaf, S. and Srour, I. (2021), "Using a data driven neural network approach to forecast building occupant complaints", *Building and Environment*, Elsevier Ltd, Vol. 200 No. March, p. 107972, doi: 10.1016/j.buildenv.2021.107972.

Baek, S., Jung, W. and Han, S.H. (2021), "A critical review of text-based research in construction: Data source, analysis method, and implications", *Automation in Construction*, Elsevier B.V., Vol. 132 No. August, p. 103915, doi: 10.1016/j.autcon.2021.103915.

Baines, P., Fill, C. and Rosengren, S. (2017), *Marketing*, Oxford University Press (ISBN: 9780198748533).

Banbury, S.P. and Berry, D.C. (2005), "Office noise and employee concentration : Identifying causes of disruption and potential improvements", *Ergonomics*, Vol. 48 No. 1, pp. 25–37, doi: 10.1080/00140130412331311390.

Bianchi, F., Nozza, D. and Hovy, D. (2021), "'FEEL-IT: Emotion and Sentiment Classification for the Italian Language'", *Proceedings of the 11th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Association for Computational Linguistics (ISBN: 9781954085183).

Borg, A. and Boldt, M. (2020), "Using VADER sentiment and SVM for predicting customer response sentiment", *Expert Systems with Applications*, Elsevier Ltd, Vol. 162, p. 113746, doi: 10.1016/j.eswa.2020.113746.

Bortolini, R. and Forcada, N. (2018), "Facility managers' perceptions on building performance assessment", *Frontiers of Engineering Management*, Vol. 5 No. 3, pp. 324–333, doi: 10.15302/J-FEM-2018010.

Bortolini, R. and Forcada, N. (2020), "Analysis of building maintenance requests using a text mining approach: building services evaluation", *Building Research and Information*, Taylor & Francis, Vol. 48 No. 2, pp. 207–217, doi: 10.1080/09613218.2019.1609291.

Brando, S., Rachinas-Lopes, P., Goulart, V.D.L.R. and Hart, L.A. (2023), "Understanding Job Satisfaction and Occupational Stressors of Distinctive Roles in Zoos and Aquariums", *Animals*, Vol. 13 No. 12, p. 2018, doi: 10.3390/ani13122018.

Bugalia, N., Tarani, V., Kedia, J. and Gadekar, H. (2022), "Machine learning-based automated classification of worker-reported safety reports in construction", *Journal of Information Technology in Construction*, Vol. 27, pp. 926–950, doi: 10.36680/j.itcon.2022.045.

Burmeister, C.P., Moskaliuk, J. and Cress, U. (2018), "Office versus leisure environments: Effects of surroundings on concentration", *Journal of Environmental Psychology*, Academic Press, Vol. 58, pp. 42–51, doi: 10.1016/J.JENVP.2018.07.011.

Cambria, E. (2016), "Affective Computing and Sentiment Analysis", *IEEE Intelligent Systems*, IEEE, Vol. 31 No. 2, pp. 102–107, doi: 10.1109/MIS.2016.31.

Cambria, E., Li, Y., Xing, F.Z., Poria, S. and Kwok, K. (2020), "SenticNet 6: Ensemble Application of Symbolic and Subsymbolic AI for Sentiment Analysis", *International Conference on Information and Knowledge Management, Proceedings*, pp. 105–114, doi: 10.1145/3340531.3412003.

Campos Fialho, B., Codinhoto, R. and Minto Fabricio, M. (2023), "Statistical critical reactive maintenance characterisation for digital twin implementation in universities", *Facilities*, No. 308379, doi: 10.1108/F-03-2023-0029.

Carvajal-Arango, D., Vásquez-Hernández, A. and Botero-Botero, L.F. (2021), "Assessment of subjective workplace well-being of construction workers: A bottom-up approach", *Journal of Building Engineering*, Vol. 36, p. 102154, doi: 10.1016/j.jobe.2021.102154.

Chanter, B. and Swallow, P. (2007), *Building Maintenance Management*, edited by Chanter, B. and Swallow, P.*Building Maintenance Management: Second Edition*, Blackwell Publishing Ltd, Oxford, UK, doi: 10.1002/9780470692011.

Chen, K., Chen, W., Li, C.T. and Cheng, J.C.P. (2019), "A BIM-based location aware AR collaborative framework for facility maintenance management", *Journal of Information Technology in Construction*, Vol. 24 No. July, pp. 360–380, doi: 10.36680/j.itcon.2019.019.

Chung, S., Moon, S., Kim, J., Kim, J., Lim, S. and Chi, S. (2023), "Comparing natural language processing (NLP) applications in construction and computer science using preferred reporting items for systematic reviews (PRISMA)", *Automation in Construction*, Elsevier B.V., Vol. 154 No. June, doi: 10.1016/j.autcon.2023.105020.

Cohen, S., Kamarck, T. and Mermelstein, R. (1983), "A global measure of perceived stress.", *Journal of Health and Social Behavior*, Vol. 24 No. 4, pp. 385–396, doi: 10.2307/2136404.

Condotta, M. and Scanagatta, C. (2023), "BIM-based method to inform operation and maintenance phases through a simplified procedure", *Journal of Building Engineering*, Vol. 65, p. 105730, doi: 10.1016/j.jobe.2022.105730.

Cotts, D.G., Roper, K.O. and Payant, R.P. (2010), *The Facility Management Handbook*, Third edit., American Management Association (ISBN: 978-0-8144-1380-7), New York, NY.

Covington, M.A. and McFall, J.D. (2010), "Cutting the Gordian knot: The moving-average type-token ratio (MATTR).", *Journal of Quantitative Linguistics*, Taylor & Francis, Covington, Michael A.: Institute for Artificial Intelligence, University of Georgia, Athens, GA, US, 30602-7415, mc@uga.edu, Vol. 17 No. 2, pp. 94–100, doi: 10.1080/09296171003643098.

D'Orazio, M., Bernardini, G. and Di Giuseppe, E. (2024), "Influence of pre-processing methods on the automatic priority prediction of native-language end-users' maintenance requests through machine learning methods", *Journal of Information Technology in Construction*, Vol. 29, doi: 10.36680/j.itcon.2024.006.

D'Orazio, M., Di Giuseppe, E. and Bernardini, G. (2022), "Automatic detection of maintenance requests: Comparison of Human Manual Annotation and Sentiment Analysis techniques", *Automation in Construction*, Elsevier B.V., Vol. 134 No. November 2021, p. 104068, doi: 10.1016/j.autcon.2021.104068.

Dam-Krogh, E.P., Rupp, R.F., Clausen, G. and Toftum, J. (2024), "Scoping review of post occupancy evaluation of office buildings with focus on indoor environmental quality and productivity", *Journal of Building Engineering*, Vol. 86, p. 108911, doi: 10.1016/j.jobe.2024.108911.

Doukari, O., Scoditti, E., Kassem, M. and Greenwood, D. (2023), "A BIM-based Techno-Economic Framework and Tool for Evaluating and Comparing Building Renovation Strategies", *Journal of Information Technology in Construction*, Vol. 28, pp. 246–265, doi: 10.36680/j.itcon.2023.012.

Ekman, P. (1999), "Basic Emotions", in Dalgleish, T. and Powers, M.J. (Eds.), *Handbook of Cognition and Emotion*, Wiley, pp. 45–60, doi: 10.1002/0470013494.ch3.

Ensafi, M., Thabet, W., Afsari, K. and Yang, E. (2023), "Challenges and gaps with user-led decision-making for prioritizing maintenance work orders", *Journal of Building Engineering*, Elsevier Ltd, Vol. 66 No. May 2022, p. 105840, doi: 10.1016/j.jobe.2023.105840.

Errandonea, I., Beltrán, S. and Arrizabalaga, S. (2020), "Digital Twin for maintenance: A literature review", *Computers in Industry*, Vol. 123, p. 103316, doi: 10.1016/j.compind.2020.103316.

Esuli, A. and Sebastiani, F. (2007), *SentiWordNet : A High-Coverage Lexical Resource*, *Technical Report - Istituto Di Scienza e Tecnologie Dell'Informazione "Alessandro Faedo" - ISTI (Available at: Http://Nmis.Isti.Cnr.It/Sebastiani/Publications/2007TR02.Pdf - Last Access 9.4.2025)*.

Gaspar, R., Pedro, C., Panagiotopoulos, P. and Seibt, B. (2016), "Beyond positive or negative: Qualitative sentiment analysis of social media reactions to unexpected stressful events", *Computers in Human Behavior*, Vol. 56, pp. 179–191, doi: 10.1016/j.chb.2015.11.040.

Gunay, H.B., Shen, W. and Newsham, G. (2019), "Data analytics to improve building performance: A critical review", *Automation in Construction*, Elsevier, Vol. 97 No. June 2018, pp. 96–109, doi: 10.1016/j.autcon.2018.10.020.

Hong, S., Kim, J. and Yang, E. (2022), "Automated Text Classification of Maintenance Data of Higher Education Buildings Using Text Mining and Machine Learning Techniques", *Journal of Architectural Engineering*, Vol. 28 No. 1, doi: 10.1061/(ASCE)AE.1943-5568.0000522.

Hu, M. and Liu, B. (2004a), "Mining and summarizing customer reviews", *KDD-2004 - Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, No. November, pp. 168–177, doi: 10.1145/1014052.1014073.

Hu, M. and Liu, B. (2004b), "Mining and summarizing customer reviews", *KDD-2004 - Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, No. August 2004, pp. 168–177, doi: 10.1145/1014052.1014073.

Hutto, C.J. and Gilbert, E. (2014), "VADER: A parsimonious rule-based model for sentiment analysis of social media text", *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014*, pp. 216–225, doi: 10.1609/icwsm.v8i1.14550.

Izobo-Martins, O., Olotuah, A., Adeyemi, E., Ayo-Vaughan, K. and Odetunmibi, O.A. (2018), "Survey dataset of building user-perceptions on the condition of public secondary school buildings in Ogun State Nigeria", *Data in Brief*, Elsevier Inc., Vol. 19, pp. 2224–2234, doi: 10.1016/j.dib.2018.06.108.

Jaafar, Syaiful Baharee. (2021), "The Relationship between Stress and Job Satisfaction", *International Journal of Business and Management*, Vol. 5 No. 1, pp. 8–12, doi: 10.26666/rmp.ijbm.2021.1.2.

Jahanger, Q.K., Louis, J., Pestana, C. and Trejo, D. (2021), "Potential positive impacts of digitalization of construction-phase information management for project owners", *Journal of Information Technology in Construction*, Vol. 26, pp. 1–22, doi: 10.36680/j.itcon.2021.001.

Joshi, A., Bhattacharyya, P. and Ahire, S. (2017), "Sentiment Resources: Lexicons and Datasets", pp. 85–106, doi: 10.1007/978-3-319-55394-8_5.

Jura, M., Spetz, J. and Liou, D.-M. (2022), "Assessing the Job Satisfaction of Registered Nurses Using Sentiment Analysis and Clustering Analysis", *Medical Care Research and Review*, Vol. 79 No. 4, pp. 585–593, doi: 10.1177/10775587211035292.

Kim, J. and Kim, N. (2022), "Quantifying Emotions in Architectural Environments Using Biometrics", *Applied Sciences (Switzerland)*, Vol. 12 No. 19, doi: 10.3390/app12199998.

Kim Wing, A.C., bin Mohammed, A.H. and bin Abdullah, M.N. (2016), "A literature review on maintenance priority - conceptual framework and directions", edited by Kamaruzzaman, S.N.B., Ali, A.S.B., Azmi, N.F.B. and Chua, S.J.L.*MATEC Web of Conferences*, Vol. 66, p. 00004, doi: 10.1051/matecconf/20166600004.

Latini, A., Di Giuseppe, E., D'Orazio, M. and Di Perna, C. (2021), "Exploring the use of immersive virtual reality to assess occupants' productivity and comfort in workplaces: An experimental study on the role of walls colour", *Energy and Buildings*, Elsevier B.V., Vol. 253, p. 111508, doi: 10.1016/j.enbuild.2021.111508.

Latini, A., Di Loreto, S., Di Giuseppe, E., D'Orazio, M. and Di Perna, C. (2023), "Crossed Effect of Acoustics on Thermal Comfort and Productivity in Workplaces: A Case Study in Virtual Reality", *Journal of Architectural Engineering*, Vol. 29 No. 2, pp. 04023009–1/10, doi: 10.1061/JAEIED.AEENG-1533.

Latini, A., Torresin, S., Oberman, T., Di Giuseppe, E., Aletta, F., Kang, J. and D'Orazio, M. (2024), "Effects of Biophilic Design interventions on university students' cognitive performance: An audio-visual experimental study in an Immersive Virtual office Environment", *Building and Environment*, Elsevier Ltd, Vol. 250 No. January, p. 111196, doi: 10.1016/j.buildenv.2024.111196.

Lei, S. and Yang, R. (2020), "Lexical richness in research articles: Corpus-based comparative study among advanced Chinese learners of English, English native beginner students and experts", *Journal of English for Academic Purposes*, Vol. 47, p. 100894, doi: 10.1016/j.jeap.2020.100894.

Li, H., Wang, Z. and Hong, T. (2021), "Occupant-Centric key performance indicators to inform building design and operations", *Journal of Building Performance Simulation*, Vol. 14 No. 6, pp. 814–842, doi: 10.1080/19401493.2021.1876771.

Liu, D., Feng, X.L., Ahmed, F., Shahid, M. and Guo, J. (2022), "Detecting and Measuring Depression on Social Media Using a Machine Learning Approach: Systematic Review", *JMIR Mental Health*, Vol. 9 No. 3, pp. 1–18, doi: 10.2196/27244.

Liu, X. and Hu, W. (2019), "Attention and sentiment of Chinese public toward green buildings based on Sina Weibo", *Sustainable Cities and Society*, Elsevier, Vol. 44 No. February 2018, pp. 550–558, doi: 10.1016/j.scs.2018.10.047.

Lu, Q., Xie, X., Kumar, A., Mary, J., Parlikad, A.K. and Schooling, J.M. (2020), "Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance", *Automation in Construction*, Elsevier, Vol. 118 No. May, p. 103277, doi: 10.1016/j.autcon.2020.103277.

Luckman, C., Wagovich, S.A., Weber, C., Brown, B., Chang, S.E., Hall, N.E. and Bernstein Ratner, N. (2020), "Lexical diversity and lexical skills in children who stutter", *Journal of Fluency Disorders*, Elsevier, Vol. 63 No. January, p. 105747, doi: 10.1016/j.jfludis.2020.105747.

Marocco, M. and Garofolo, I. (2021), "Operational text-mining methods for enhancing building maintenance management", *Building Research and Information*, Vol. 49 No. 8, pp. 893–911, doi: 10.1080/09613218.2021.1953368.

Martinis, M.C., Zucco, C. and Cannataro, M. (2022), "An Italian lexicon-based sentiment analysis approach for medical applications", *Proceedings of the 13th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, BCB 2022*, doi: 10.1145/3535508.3545594.

Marzouk, M. and Enaba, M. (2019), "Text analytics to analyze and monitor construction project contract and correspondence", *Automation in Construction*, Elsevier, Vol. 98 No. December 2017, pp. 265–274, doi: 10.1016/j.autcon.2018.11.018.

Maslesa, E. and Jensen, P.A. (2019), "Managing environmental building performance through IT systems", *Facilities*, Vol. 38 No. 3/4, pp. 235–251, doi: 10.1108/F-01-2019-0013.

McCarthy, P.M. and Jarvis, S. (2010), "MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment", *Behavior Research Methods*, Vol. 42 No. 2, pp. 381–392, doi: 10.3758/BRM.42.2.381.

Mo, Y., Zhao, D., Du, J., Syal, M., Aziz, A. and Li, H. (2020), "Automated staff assignment for building maintenance using natural language processing", *Automation in Construction*, Elsevier, Vol. 113 No. February, p. 103150, doi: 10.1016/j.autcon.2020.103150.

Mucci, N., Giorgi, G., Cupelli, V., Gioffrè, P.A., Rosati, M.V., Tomei, F., Tomei, G., *et al.* (2015), "Work-related stress assessment in a population of Italian workers. The Stress Questionnaire", *Science of The Total Environment*, Vol. 502, pp. 673–679, doi: 10.1016/j.scitotenv.2014.09.069.

Nandwani, P. and Verma, R. (2021), "A review on sentiment analysis and emotion detection from text", *Social Network Analysis and Mining*, Springer Vienna, Vol. 11 No. 1, p. 81, doi: 10.1007/s13278-021-00776-6.

Nielsen, F.Å. (2011), "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs", *CEUR Workshop Proceedings*, Vol. 718, (ISSN: 16130073), pp. 93–98.

Nijhawan, T., Attigeri, G. and Ananthakrishna, T. (2022), "Stress detection using natural language processing and machine learning over social interactions", *Journal of Big Data*, Vol. 9 No. 1, doi: 10.1186/s40537-022-00575-6.

Nolé Fajardo, M.L., Higuera-Trujillo, J.L. and Llinares, C. (2023), "Lighting, colour and geometry: Which has the greatest influence on students' cognitive processes?", *Frontiers of Architectural Research*, Vol. 12 No. 4, pp. 575–586, doi: 10.1016/j.foar.2023.02.003.

O'Brien, W., Gaetani, I., Carlucci, S., Hoes, P.-J. and Hensen, J.L.M. (2017), "On occupant-centric building performance metrics", *Building and Environment*, Pergamon, Vol. 122, pp. 373–385, doi: 10.1016/j.buildenv.2017.06.028.

O'Neill, M.J. and Carayon, P. (1993), "The Relationship between Privacy, Control, and Stress Responses in Office Workers", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 37 No. 7, pp. 479–483, doi: 10.1177/154193129303700702.

Pan, Y. and Zhang, L. (2021), "Roles of artificial intelligence in construction engineering and management: A critical review and future trends", *Automation in Construction*, Elsevier B.V., Vol. 122 No. October 2020, p. 103517, doi: 10.1016/j.autcon.2020.103517.

Parisi, F., Fanti, M.P. and Mangini, A.M. (2021), "Information and Communication Technologies applied to intelligent buildings: a review", *Journal of Information Technology in Construction*, Vol. 26, pp. 458–488, doi: 10.36680/j.itcon.2021.025.

Pärn, E.A., Edwards, D.J. and Sing, M.C.P. (2017), "The building information modelling trajectory in facilities management: A review", *Automation in Construction*, Elsevier B.V., Vol. 75, pp. 45–55, doi: 10.1016/j.autcon.2016.12.003.

Peter, N.., Ibisola, A.S., Oluwatayo, A.A., Omotosho, V.J. and Okagbue, H.I. (2023), "Users' perception of the physical work environment in Covenant University", *Materials Today: Proceedings*, Elsevier Ltd, No. December 2022, doi: 10.1016/j.matpr.2023.08.079.

Pishdad-Bozorgi, P., Gao, X., Eastman, C. and Self, A.P. (2018), "Planning and developing facility management-enabled building information model (FM-enabled BIM)", *Automation in Construction*, Elsevier, Vol. 87 No. December 2017, pp. 22–38, doi: 10.1016/j.autcon.2017.12.004.

Presti, P., Ruzzon, D., Avanzini, P., Caruana, F., Rizzolatti, G. and Vecchiato, G. (2022), "Measuring arousal and valence generated by the dynamic experience of architectural forms in virtual environments", *Scientific Reports*, Nature Publishing Group UK, pp. 1–13, doi: 10.1038/s41598-022-17689-9.

Rampini, L. and Cecconi, F.R. (2022), "Artificial intelligence in construction asset management: a review of present status, challenges and future opportunities", *Journal of Information Technology in Construction*, Vol. 27, pp. 884–913, doi: 10.36680/j.itcon.2022.043.

Rasheed, E.O., Khoshbakht, M. and Baird, G. (2021), "Time spent in the office and workers' productivity, comfort and health: A perception study", *Building and Environment*, Elsevier Ltd, Vol. 195 No. February, p. 107747, doi: 10.1016/j.buildenv.2021.107747.

Russo, A. and Coco, D. (2023), "Quantify how space mission influence geopolitical dynamics? A security and social policy approach", *ArXiv Preprint ArXiv:2301.03538*, pp. 1–18.

Saleem, S., Prasad, R., Vitaladevuni, S., Pacula, M., Crystal, M., Marx, B., Sloan, D., *et al.* (2012), "Automatic detection of psychological distress indicators and severity assessment from online forum posts", *24th International Conference on Computational Linguistics - Proceedings of COLING 2012: Technical Papers*, Vol. 5, (available at aclanthology.org/C12-1145.pdf, lasta access: 09/04/2025), pp. 2375–2388.

Sánchez-Garrido, A.J., Navarro, I.J., García, J. and Yepes, V. (2023), "A systematic literature review on modern methods of construction in building: An integrated approach using machine learning", *Journal of Building Engineering*, Vol. 73, p. 106725, doi: 10.1016/j.jobe.2023.106725.

Sánchez-Rada, J.F. and Iglesias, C.A. (2019), "Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison", *Information Fusion*, Vol. 52 No. May, pp. 344–356, doi: 10.1016/j.inffus.2019.05.003.

Sanni-Anibire, M.O., Zin, R.M. and Olatunji, S.O. (2021), "Machine learning - based framework for construction delay mitigation", *Journal of Information Technology in Construction*, Vol. 26, pp. 303–318, doi: 10.36680/j.itcon.2021.017.

Sanzana, M.R., Maul, T., Wong, J.Y., Abdulrazic, M.O.M. and Yip, C.C. (2022), "Application of deep learning in facility management and maintenance for heating, ventilation, and air conditioning", *Automation in Construction*, Elsevier B.V., Vol. 141 No. February, p. 104445, doi: 10.1016/j.autcon.2022.104445.

Schmidt, T., Burghardt, M. and Dennerlein, K. (2018), "Sentiment annotation of historic German plays: An empirical study on annotation behavior", *CEUR Workshop Proceedings*, Vol. 2155, (ISSN: 16130073), pp. 47–52.

Shalabi, F. and Turkan, Y. (2020), "BIM–energy simulation approach for detecting building spaces with faults and problematic behavior", *Journal of Information Technology in Construction*, Vol. 25, pp. 342–360, doi: 10.36680/j.itcon.2020.020.

Shamshiri, A., Ryu, K.R. and Park, J.Y. (2024), "Text mining and natural language processing in construction", *Automation in Construction*, Vol. 158, p. 105200, doi: https://doi.org/10.1016/j.autcon.2023.105200.

Valença, J., Morin, K., Jouen, N., Olivo, N., Torres-Gonzalez, M., Mendes, M.P. and Silva, A. (2024), "Feeling-BIM: A digital model to support maintenance decisions, based on automatic inspection and dwellers' feelings", *Journal of Building Engineering*, Vol. 87, p. 108937, doi: 10.1016/j.jobe.2024.108937.

Valiyappurakkal, V.K. (2022), "Evaluation of the effect of preconception due to past experience on the perception of building performance of an Indian hostel", *Building and Environment*, Elsevier Ltd, Vol. 215 No. November 2021, p. 108951, doi: 10.1016/j.buildenv.2022.108951.

Vandierendonck, A., Liefooghe, B. and Verbruggen, F. (2010), "Task Switching: Interplay of Reconfiguration and Interference Control", *Psychological Bulletin*, Vol. 136 No. 4, pp. 601–626, doi: 10.1037/a0019791.

Vizer, L.M., Zhou, L. and Sears, A. (2009), "Automated stress detection using keystroke and linguistic features: An exploratory study", *International Journal of Human-Computer Studies*, Elsevier, Vol. 67 No. 10, pp. 870–886, doi: 10.1016/j.ijhcs.2009.07.005.

Wang, Z., Ho, S.B. and Cambria, E. (2020), "A review of emotion sensing: categorization models and algorithms", *Multimedia Tools and Applications*, Multimedia Tools and Applications, Vol. 79 No. 47–48, pp. 35553–35582, doi: 10.1007/s11042-019-08328-z.

Wong, J.K.W., Ge, J. and He, S.X. (2018), "Digitisation in facilities management: A literature review and future research directions", *Automation in Construction*, Elsevier, Vol. 92 No. May, pp. 312–326, doi: 10.1016/j.autcon.2018.04.006.

Woods, K., Hashimoto, B. and Brown, E.K. (2023), "A multi-measure approach for lexical diversity in writing assessments: Considerations in measurement and timing", *Assessing Writing*, Elsevier Inc., Vol. 55 No. November 2022, p. 100688, doi: 10.1016/j.asw.2022.100688.

Yin, J., Zhu, S., MacNaughton, P., Allen, J.G. and Spengler, J.D. (2018), "Physiological and cognitive performance of exposure to biophilic indoor environment", *Building and Environment*, Elsevier, Vol. 132 No. January, pp. 255–262, doi: 10.1016/j.buildenv.2018.01.006.

Yousefli, Z., Nasiri, F. and Moselhi, O. (2020), "Maintenance workflow management in hospitals: An automated multi-agent facility management system", *Journal of Building Engineering*, Elsevier Ltd, Vol. 32 No. April, p. 101431, doi: 10.1016/j.jobe.2020.101431.

Zalejska-Jonsson, A. (2020), "Does facility management affect perception of building quality?", *Facilities*, Vol. 38 No. 7/8, pp. 559–576, doi: 10.1108/F-02-2019-0026.

Zenker, F. and Kyle, K. (2021a), "Investigating minimum text lengths for lexical diversity indices", *Assessing Writing*, Vol. 47, p. 100505, doi: 10.1016/j.asw.2020.100505.

Zenker, F. and Kyle, K. (2021b), "Investigating minimum text lengths for lexical diversity indices", *Assessing Writing*, Vol. 47, p. 100505, doi: 10.1016/j.asw.2020.100505.

Zhang, C., Yu, M.C. and Marin, S. (2021), "Exploring Public Sentiment on Enforced Remote Work During COVID-19", *Journal of Applied Psychology*, Vol. 106 No. 6, pp. 797–810, doi: 10.1037/apl0000933.

Zhang, T., Schoene, A.M., Ji, S. and Ananiadou, S. (2022), "Natural language processing applied to mental illness detection: a narrative review", *Npj Digital Medicine*, Springer US, Vol. 5 No. 1, pp. 1–13, doi: 10.1038/s41746-022-00589-7.

Zhang, Y., Chen, J., Liu, H., Chen, Y., Xiao, B. and Li, H. (2024), "Recent advancements of human-centered design in building engineering: A comprehensive review", *Journal of Building Engineering*, Vol. 84, p. 108529, doi: 10.1016/j.jobe.2024.108529.

Zhou, L., Burgoon, J.K., Nunamaker, J.F. and Twitchell, D. (2004), "Automating Linguistics-Based Cues for detecting deception in text-based asynchronous computer-mediated communication", *Group Decision and Negotiation*, Vol. 13 No. 1, pp. 81–106, doi: 10.1023/B:GRUP.0000011944.62889.6f.

Zhu, M., Yang, Y. and Hsee, C.K. (2018), "The Mere Urgency Effect", edited by Johar, G. and Lee, L.*Journal of Consumer Research*, doi: 10.1093/jcr/ucy008.

Zubair, M.U., Zhang, X. and Hassan, M.U. (2024), "Aligning expectations and perceptions of elevator maintenance stakeholders: an application of expectation confirmation theory", *Building Research & Information*, Vol. 52 No. 4, pp. 418–430, doi: 10.1080/09613218.2023.2236252.

APPENDIX A: NOMENCLATURE

ACRONYMS AND SYMBOLS

CMMS	computer management maintenance systems
ML	machine learning
NLP	natural language processing
GFA	gross floor area
REGEX	regular expression
TIME-TO-SOLVE	time necessary to solve a maintenance request
WM	indicators on workplace maintenance needs
JB	indicators on workplace job-related stress indicators of both end-users (i.e. organization workers) and maintenance technicians
PQ	indicators on end-users and technicians' perception of maintenance quality in the workplace
LD	lexical diversity
TTR	Type-To-Token ratio
MATTR	Moving Average TTR
VCS	VADER compound score
MR	Number of maintenance requests

SUBSCRIPTS

j	building complex
k	building occupant or group/subgroup of occupants within end-users (i.e. organization workers) or maintenance technicians
EU	end-users of the building stock, i.e. organization workers
TE, TECH	maintenance technicians working in the building stock
Ttot	sum of total time to solve all maintenance requests
Stat	average (in case of normal data distribution) or median value (in case of non-normal data distribution and to limit the impact of outliers)
T	significant statistical distribution value of the time necessary to solve the requests (TIME-TO-SOLVE_stat)
MATTRtot	overall MATTR-based stress detection related to the maintenance requests
MATTR	significant statistical distribution value of the MATTR values (MATTR_stat)
Vtot	sum of VADER compound score-based assessment of each request
V	significant statistical distribution value of the VADER compound score-based assessment of requests (VCS_stat)

APPENDIX B

The FEEL-IT Italian model performs emotion and sentiment analysis on Italian texts. The authors (Bianchi, 2021) evaluated the sentiment model's performance (accuracy: 0.81) using the SENTIPOLC16 dataset from Evalita, by collapsing the original FEEL-IT emotion classes into two: grouping *"joy"* into the positive class, and *"anger"*, *"fear"*, and *"sadness"* into the negative class.

VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon- and rule-based sentiment analysis tool. VADERITA is a specialized lexicon adapted for the Italian language (Hutto, 2014). VADER expresses sentiment through a *compound score*, computed by summing the valence scores of each word in the lexicon, adjusted according to predefined rules, and then normalized to a range between -1 (most extreme negative) and +1 (most extreme positive).

Although VADER and FEEL-IT results are not fully comparable (e.g., *sadness* is not always interpreted as a negative emotion), we compared the results obtained by the two methods on all sentences. Positive VADER compound scores (> 0) were matched with FEEL-IT-labelled *positive* sentiments, while non-positive compound scores (≤ 0) were matched with FEEL-IT-labelled *negative* sentiments.

Table 1 shows the comparison results. When both VADER and FEEL-IT indicate either a positive or a negative sentiment, we consider this a *positive correspondence* (condition = true). Conversely, when one method indicates a positive sentiment and the other a negative one, we consider it a *negative correspondence*. The data reveals the slight prevalence of positive correspondence (55%). This is due to the high number of sentences labelled with the sadness emotion by FEEL-IT.

Table 3: Count of positive (true) or negative (false) correspondence between the two methods.

Condition	**Count**
True	11715
False	9509
Total	**21224**