Using Demographic Features for the Prediction of Basic Human Values Underlying Stakeholder Motivation

Adam Szekeres, Pankaj Shivdayal Wasnik and Einar Arthur Snekkenes
Department of Information Security and Communication Technology, Norwegian University of Science and Technology,

Keywords: Information Security Risk Management, Stakeholder Motivation, Psychological Perspective, Motivational Profiles.

Abstract: Human behavior plays a significant role within the domain of information security. The Conflicting Incentives Risk Analysis (CIRA) method focuses on stakeholder motivation to analyze risks resulting from the actions of key decision makers. In order to enhance the real-world applicability of the method, it is necessary to characterize relevant stakeholders by their motivational profile, without relying on direct psychological assessment methods. Thus, the main objective of this study was to assess the utility of demographic features-that are observable in any context-for deriving stakeholder motivational profiles. To this end, this study utilized the European Social Survey, which is a high-quality international database, and is comprised of representative samples from 23 European countries. The predictive performances of a pattern-matching algorithm and a machine-learning method are compared to establish the findings. Our results show that demographic features are marginally useful for predicting stakeholder motivational profiles. These findings can be utilized in settings where interaction between a stakeholder and an analyst is limited, and the results provide a solid benchmark baseline for other methods, which focus on different classes of observable features for predicting stakeholder motivational profiles.

1 INTRODUCTION

Information security is considered to be a highly technical domain, where research on the human element gets relatively low attention, given the involvement and impact of individuals on the system's safety and security. However, "...people are responsible for stealing passwords, committing intellectual property crimes, skimming financial accounts, selling information to competitors, breaking into databases, cybersnooping, and committing a host of other offenses against organizations and their systems. Ironically, the disciplines that assess, evaluate, and solve human based problems have not been an integral part of the information security measures used to protect data..." (Gudaitis, 1998). It is suggested that there is a need for synthesis between various disciplines in order to improve on the attempts that aim to protect against threats to information systems. More than a decade later, Greitzer and Hohimer (Greitzer and Hohimer, 2011) concluded that insider threats ranked among the most problematic cyber-security challenges that threaten government and industry information infrastructures. Furthermore, they identified that there were no systematic methods that provided a complete and effective approach to preventing undesirable actions (e.g. data leakage, espionage, and sabotage).

More recent incidents (e.g. using technical expertise and insider privileges to reprogram Smart Meters (Krebs, 2012), cheating with emission rates (Arora, 2017), financial misreporting (Kulik et al., 2008), creating abusive websites (Franklin, 2014), etc.) also call for methods that incorporate intentional, deliberate human behavior into risk assessments. While the specific details of the enumerated incidents vary greatly, they are still united by some common features:

- It is possible to identify a person or a group who had a strong motivation to take certain actions.
- It is possible to identify a person or a group who suffered the consequences of those actions but who were unintentionally exposed to those transactions.

Such situations are recognized in the economic literature as negative externalities (Liebowitz and Margolis, 1994) and the concept has been applied within the domain of information security, where motivated actors have the potential to exert a negative influence on

Szekeres, A., Wasnik, P. and Snekkenes, E.
Using Demographic Features for the Prediction of Basic Human Values Underlying Stakeholder Motivation.
DOI: 10.5220/0007694203770389
In *Proceedings of the 21st International Conference on Enterprise Information Systems (ICEIS 2019)*, pages 377-389
ISBN: 978-989-758-372-8

a large number of other stakeholders who have little influence on the outcome of those actions (Anderson and Moore, 2009).

Assessing stakeholder motivation could be the key to preparing against such events, since motivation is a central concept in understanding human behavior; it aims to answer the question concerning why people do the things they do (Forbes, 2011). During the past centuries, researchers have generated a vast number of theoretical constructs and systems which vary in the level of the analysis (e.g. instincts, biologically determined drives, needs, social and cognitive motivations), the scope (e.g. general principles vs. task-specific motivations), and the terminology. Through describing stakeholder motivation we can enable the prediction of future behaviors and check whether the likely behavior is in alignment with the goals of other affected stakeholders. However, people are not expected to cooperate in any analysis that aims to assess their motivations for risk-analysis purposes. Therefore, the main goal of the present study is to contribute to the information security risk management literature by investigating the utility of demographic features for deriving stakeholder motivational profiles in contexts where no direct interaction between the subject and analyst is assumed.

Following the Problem Statement and Research Questions, Section 2 describes the risk analysis method under development, and it's connection to the theory of basic human values. Section 3 explains how a publicly available high-quality dataset was utilized in the study, which is followed by describing the results in Section 4. Section 5 provides an overview of the conducted work, and Section 6 concludes with directions for future work.

1.1 Problem Statement

The main objective of this work is to investigate how stakeholder motivation can be predicted by utilizing publicly observable individual characteristics (e.g. demographic variables). The end goal is the development of a predictive model that can be utilized by an observer to derive the motivational profile of a previously unknown subject by collecting and aggregating various forms of publicly observable features connected to the subject.

1.2 Research Questions

To address the problem statement, the following research questions have been formulated:

1. To what extent can demographic features be utilized to construct stakeholder motivational profiles?
2. How well do different predictive models perform in terms of inferring stakeholder motivational profiles?

2 RELATED WORK

This section provides an overview of the risk-analysis method under development, the motivational theory, and the related constructs that were included in the study.

2.1 Conflicting Incentives Risk Analysis

The importance of understanding stakeholder motivation is emphasized within the Conflicting Incentives Risk Analysis (CIRA) method (Rajbhandari and Snekkenes, 2013). This method identifies the stakeholders (i.e. individuals), the actions that can be taken by the stakeholders, as well as the consequences of these actions. A stakeholder is a physical person who has some interest in the outcomes of his actions. The procedure identifies two types of stakeholders: the *Strategy owner* (the person who is capable of executing an action) and the *Risk owner* (whose perspective is taken-the person at risk). Each stakeholder's motivation is modeled on the concept of utility, which entails the consideration of the benefit of the action performed from the perspective of the stakeholder. This cumulative utility encompasses several utility factors, each representing aspects of life considered important by the corresponding stakeholders. Two types of risks are identified in the method: Threat risk refers to the perceived decrease in the total utility of the risk owner and Opportunity Risk refers to the lack of potential increase in utility because the strategy owner is not motivated enough to take actions that would be beneficial for the Risk owner. Therefore, risk is conceptualized as a misalignment of incentives between these two classes of stakeholders, and risk identification is about uncovering activities that would be beneficial for the Strategy owner, and potentially harmful for the Risk owner, or vice versa (Snekkenes, 2013). Therefore, Threat risk closely resembles the concept of moral hazard; it captures a wide range of behaviors that are beneficial for one party and detrimental for another (i.e. the strategy owner inflicting negative externalities on the risk owner) (Dembe and Boden, 2000). Previous work explored the feasibility of inferring key stakeholders' motivational profiles based on the linguistic analysis of interviews given by inaccessible subjects (Szekeres and Snekkenes, 2018).

2.2 Theory of Basic Human Values

The theory of basic human values, developed by Schwartz, (Schwartz, 1994) identifies ten distinct values that are universally recognized across various cultures, and it provides a unified and comprehensive view on human motivation. The theory incorporates several previous approaches that emphasized the centrality of values in human behavior (e.g. Hofstede and Rokeach on cultural differences (Schwartz, 1992)). Values both represent desirable end-goals and prescribe desirable ways of acting. Schwartz summarizes the six core features that characterize values:

- "Values are beliefs linked to affect.
- Values refer to desirable goals that motivate actions.
- Values transcend specific actions and situations.
- Values serve as standards or criteria.
- Values are ordered by importance.
- The relative importance of multiple values guide actions."

Furthermore, all of the ten distinct values in the theory encapsulate one of the three key motivational aspects that are grounded in the universal requirements of human existence: the needs of individuals as biological organisms, the requisites of coordinated social interaction, and the survival and welfare needs of groups. Values guide behavior, given that the decision context, or situation activates the relevant values. The ten values form a circular structure that captures a motivational continuum, where adjacent values are compatible with each other, while opposing values are in conflict. The ten values are grouped under four higher dimensions, as represented by Figure 1 (Schwartz, 2012).

Goldberg, Sweeney, Merenda, and Hughes (Goldberg et al., 1998) describe how one of the most enduring topics in the history of psychometrics is the strength of association between group and individual differences, and the many controversies centered around the issue of how various demographically defined groups differ in terms of important human attributes. In their study, they investigated the differences between the Big Five personality traits and four demographic variables (i.e. gender, age, education, and ethnic status). The study concluded that most demographic-personality associations are of trivial size, with an average correlation of 0.08 (across the four demographic variables and the five personality dimensions included in the study). However, these results are not directly comparable to the value-demographic association yet, they nevertheless pro-

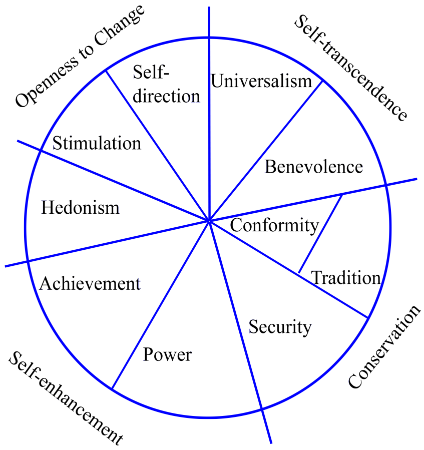

Figure 1: Circular value structure, with 4 higher dimensions comprising of the 10 basic human values.

vided some initial insights into the strength of associations between demographic features and psychological variables. Schwartz (Schwartz, 2007) discusses the reciprocal relationship between value priorities and life circumstances and provides empirical evidence on the hypothetical relationships. Choices guided by values influence the life circumstances, but certain life circumstances (e.g. the type of profession, raising children, etc.) also affect the possibility of, and constraints placed upon, enacting particular choices. People tend to adapt their values to fit into their life circumstances by upgrading the importance of values that are readily attainable, while downgrading the importance of values of which the pursuit is blocked. As people's demographic variables (e.g. age, gender, education, income level, etc.) largely impact the circumstances to which they are exposed, these differences are expected to have a direct effect on the value priorities. Based on the value system's structure, the following subsections present validated and hypothesized relationships between demographic variables and value priorities based on (Schwartz, 2007).

2.2.1 Age

Due to the general decline of physical strength and cognitive abilities, aging is expected to increase the importance of Security values, as the capacity to deal with change declines. Therefore, the opposing Stimulation value might decrease in importance as novelty and risk is viewed as increasingly threatening. Conformity and Tradition values might increase in importance, while Hedonism could potentially decrease due to the dulling of the senses. Achievement and Power values may also decrease in importance since older people become less able to perform demanding tasks and obtain social approval.

2.2.2 Life Stages

In early adulthood people are primarily concerned with establishing themselves within the domains of work and family. The pursuit of Achievement and Stimulation values comes at the expense of the Security, Conformity, and Tradition values. Later, the motivation shifts to preserving the status already attained, both in the professional and in the family domains. The possibility of radical change narrows and responsibilities constrain the opportunities for risk-taking. Taking these factors into consideration, it is expected that people in their middle adulthood express a stronger preference for values encompassed in the Conservation category. At later stages, close to retirement, the opportunities for expressing Achievement, Power, Stimulation, and Hedonism values further decrease.

2.2.3 Gender

In a cross-cultural, large scale study, Schwartz and Rubel investigated gender differences in value priorities (Schwartz and Rubel, 2005). The findings suggest that men attribute more importance to Self-enhancement and Openness to change values than women do, while for Self-transcendence values, the reverse is true. The differences are generally small, and account for less variance than age and culture do, for example.

2.2.4 Education

An explanation for the association between the level of education and the values is offered in (Schwartz, 2007). According to the hypothesis education requires intellectual openness, and flexibility that is associated with Self-direction values. Challenging existing views and norms can be linked to a lower importance assigned to Conservation values, as they promote conformity and tradition. Furthermore, there might be a positive correlation with Achievement values as performance and meeting external standards is increasingly important as the level of education rises.

2.2.5 Country

The challenges faced by nations in organizing human activities are similar, but nations differ in the importance they attribute to certain values (Schwartz, 2013). When values are analyzed at the societal level, three bipolar dimensions can be identified based on the alternative resolutions to each of the problems affecting all societies: Embeddedness vs. Autonomy (affective and intellectual), Hierarchy vs. Egalitarianism, and Mastery vs. Harmony. The importance assigned by various countries to the previous dimensions gives rise to eight distinct cultural regions, representing vague differences among cultures: Western Europe, East-Central Europe, Eastern Europe, Latin America, English-Speaking, Confucian, South-East Asia, and Africa-Middle East.

2.2.6 Occupation

Another study by Knafo and Sagiv (Knafo and Sagiv, 2004) investigated the relationship between values and occupational choices. The survey-based study showed that the 32 occupations under investigation clustered according to the motivational profiles of the individuals within the profession, and that these clusters fit well into Holland's work typology. Universalism values negatively correlated with the Enterprising work environment, while Social environments correlated positively with both Universalism and Benevolence values, and correlated negatively with power and Achievement values. Artistic work environments correlated negatively with Conformity values while the Investigative environments correlated positively with Openness to change values.

These results suggest that there are meaningful and detectable differences among various groups of people. However, to our knowledge, there is no existing study that investigates how well the motivational profile can be predicted when solely based upon demographic features. Therefore, this study aims to establish predictive models from a high-quality database that contains representative samples from 23 European countries.

3 MATERIALS AND METHODS

3.1 Sample and Procedure

The European Social Survey (ESS), round 8, edition 2.0, (N.A., 2018) served as the main source of answers to the research questions. The high-quality cumulative dataset contains individual-level data from 23 countries (Austria, Belgium, the Czech Republic, Estonia, Finland, France, Germany, Hungary, Iceland, Ireland, Israel, Italy, Lithuania, the Netherlands, Norway, Poland, Portugal, the Russian Federation, Slovenia, Spain, Sweden, Switzerland, and the United Kingdom), gathered using strict probability sampling methods. The survey's main objectives are to monitor and interpret changing public attitudes in Europe, to investigate relevant societal issues, and to establish social indicators across Europe. The original dataset contains a total of (n = 44 387) individual respondents with 536 variables. The ESS has been conducted every two years since 2001 across European many countries. The survey consists of two main parts:

- The core module - covers a wide range of topics (e.g. politics, social trust, household, socio-demographics, human values, etc.) that largely remain the same in each round to allow for longitudinal observations.
- The rotating module - increases the scope of the survey by focusing on specific topics between different times of administration (e.g. immigration, economic morality, justice, democracy, climate change, etc.)

3.2 Measures

In order to address the research questions, the following preparation procedures were conducted on the original cumulative dataset. In the first step, the complete list of variables (N_{vars} = 536) was screened and then it was sorted into four main categories (demographics, attitudes, behaviors, and others). The next step focused on identifying the demographic attributes that met the inclusion criteria (i.e. the predictor variables should be publicly observable and easily identifiable by an observer). This resulted in a list of demographic variables being included in the present analysis (N_{vars} = 14), accompanying the basic human values. Table 1 contains the list of independent variables selected for the analysis. We aimed at maximizing the number of subjects with valid responses, therefore, the next step was to investigate the number of missing values in the sample. Since our objective was to analyze the predictability of the motivational profiles of individuals who are actively employed we used a listwise deletion of subjects with missing values on any of the remaining variables. The listwise removal of data is justified by the fact that most of the missing data was attributed to four variables associated with employment relations (the last four variables in Table 1), with a not-applicable label (e.g. the not actively working age-group) which contributed to a total of 7255 subjects with missing data, while the remaining missing data (n = 385) was distributed among the ten other independent variables (with the labels: refusal, do not know, no answer, not available). While it was not possible to determine whether the data was missing at random, completely at random, or not at random for the remaining small number of cases, the relatively small number enabled deletion without introducing a bias into the models. Additionally, the 89 levels of variable "Type of industry working for" were grouped according to the NACE rev. 2. section codes, resulting in 21 higher level groups (Eurostat, 2008) providing larger groups within occupational categories. The ESS dataset contains raw responses for the Human Values Scale, which is a 21-item survey instrument designed for self-assessment. In order to compute ground-truth scores from the raw item-level responses, we followed the procedures described in the accompanying manual (Schwartz, 2016). Finally, all dependent variables (the ten basic values) were normalized to a range of [0-1] through the following method: $X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$, since it is provides a linear transformation and keeps the relationships among the original data (Patro and Sahu, 2015).

Table 1: List of observable features used as predictors.

	Categorical variable (Yes/No)	Number of categories
Country	Y	23
Gender	Y	2
Age	N	-
Domicile	Y	5
Belonging to religion	Y	2
Belonging to a minority ethnic group	Y	2
Number of people living in the same household	N	-
Living with partner	Y	2
Ever had a divorce	Y	2
Highest level of education	N	-
Employment relation	Y	3
Supervising others at work	Y	2
Type of industry working in (NACE rev.2)	Y	21
Type of organization working for	Y	6

4 RESULTS

This section describes the experiments conducted on the ESS dataset and the results obtained from two different types of analytic techniques. All subjects with valid responses on the 14 features were included in the final analyses (n = 36 747): 48.5% of the subjects were males and the mean age of all respondents was

50.41 years (SD = 17.55). Furthermore, the database was randomized and divided into three sets:

- Training set: 60%
- Development set: 20%
- Testing set: 20%

4.1 Multiple Linear Regression Approach

Several multiple linear regressions (LRs) were conducted to identify the most suitable set of features that can be utilized for predicting the human value scores based on the observable features presented in Table 1. This part of the analysis was conducted using IBM SPSS 25's automatic linear modeling module, which includes supervised merging of the categories, outlier detection, and several feature-selection methods (Yang, 2013). For each of the ten basic values, the first step involved the assessment of the maximum possible predictive accuracy by using all the features, which aided us in providing an estimate of the highest potential accuracy achievable. Next, predictors were entered into the models using the forward stepwise selection algorithm. At each step, variables not yet included in the model were tested for inclusion until no variables met the inclusion criteria, using a limit of 4 as the maximum number of effects in the final model. This reflects a decision to trade-off a marginal improvement in accuracy for a simpler model with lower costs in terms of data collection. The procedure resulted in two models for each of the ten values, as shown in Table 2. Performance was measured by the R^2 (coefficient of determination), ranging between 0-1, which is a well-established, common measure of the success of predicting the dependent variable from the independent variables (Nagelkerke et al., 1991). Formula: $R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$, where SS_{res} is the sum of the residual squares and SS_{tot} is the total sum of squares. This procedure enabled us to assess the observable feature's utility in terms of predicting the ten basic values, and to identify an optimal set of features that can sufficiently cover all the basic human values considering the added utility of each feature relative to what is already included in the model.

Table 2: Statistics of R^2 values for the Linear Regression approach. In the last column, values in parentheses represent the number of features used in the final model.

	Max possible R^2	Final R^2
Achievement	0.23	0.16 (2)
Benevolence	0.22	0.16 (2)
Conformity	0.17	0.11 (2)
Hedonism	0.22	0.18 (2)
Power	0.24	0.18 (1)
Security	0.20	0.12 (3)
Self-Direction	0.16	0.09 (3)
Stimulation	0.16	0.09 (2)
Tradition	0.24	0.14 (4)
Universalism	0.18	0.13 (3)

Figure 2 presents each dependent variable with the best set of demographic variables, that account for the largest amount of explained variance (see the 'Final R^2' column from Table 2 for the corresponding models). The colored bars represent demographic features that were included in the final models and their length represents the amount of variance explained by the corresponding variable. The white bars represent the amount of unexplained variance for each value, and as such, they express the amount of remaining uncertainty regarding a subject's motivational profile. Figure 4 and Figure 5 in the Appendix provides the de-

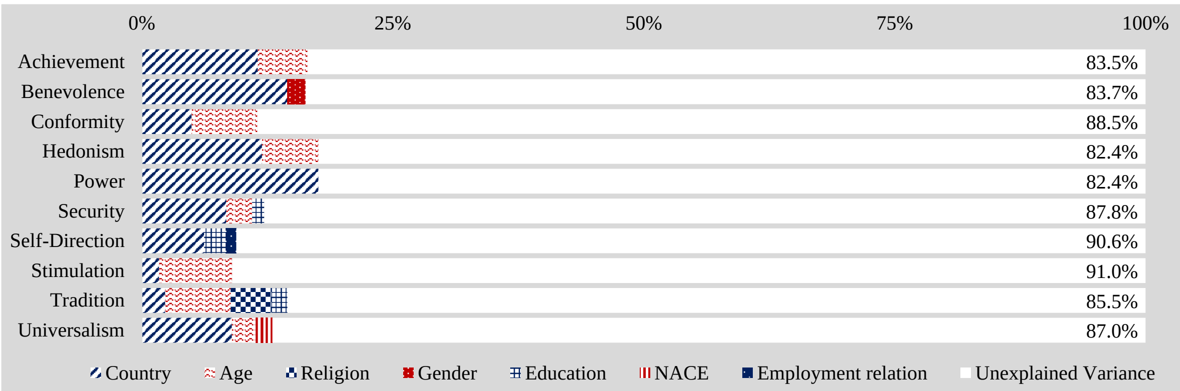

Figure 2: Feature importance for predicting the 10 basic human values from observable features by the LR approach relative to unexplained variance expressed in terms of R^2 scores.

tails of all the final regression models for each of the ten values.

4.2 Machine Learning Approach

This experiment utilized a machine learning (ML) approach for the prediction of the same set of basic human values. The regression models were trained using the H_2O.ai API, which is an open-source ML platform (H2O.ai, 2018b). The Distributed Random Forest (DRF) regression algorithm was chosen for building models for each of the ten values separately, since the algorithm can properly handle categorical variables with several levels (H2O.ai, 2018a), and also provides useful internal estimates of error, correlation, and variable importance metrics (Breiman, 2001). Furthermore, when given a training dataset, the DRF creates a forest of classification (or regression trees) instead of a single tree.

4.2.1 DRF Training

During the training stage, the models were trained using a 5-fold cross validation procedure to obtain the final model of the training set. Table 3 presents the mean and the standard deviation of the root-mean square error (RMSE) scores for all of the five folds.

Table 3: Mean and SD of RMSE and R^2 for 5 fold cross validation training.

Dependent Variable	RMSE		R^2	
	Mean	SD	Mean	SD
Achievement	0.128	0.0002	0.141	0.0090
Benevolence	0.098	0.0009	0.126	0.0096
Conformity	0.127	0.0005	0.097	0.0041
Hedonism	0.106	0.0005	0.139	0.0134
Power	0.120	0.0004	0.159	0.0033
Security	0.112	0.0004	0.109	0.0095
Self-Direction	0.113	0.0013	0.072	0.0034
Stimulation	0.114	0.0008	0.074	0.0064
Tradition	0.104	0.0006	0.122	0.0092
Universalism	0.102	0.0008	0.106	0.0070

The RMSE scores indicate the absolute fit of the model as it is the square root of the variance of the residuals in the prediction model. As such it is a good measure of the model's predictive accuracy. The RMSE can be interpreted as the standard deviation of the unexplained variance and it has the same unit as the dependent variable (Grace-Martin, 2008). The models were tuned on the hyperparameter 'number of trees' using the development set. The hyperparameter tuning favoured a higher number of trees. However, increasing the number of trees beyond 50 did not result in a significant improvement in terms of the RMSE. Therefore, for all of the ten models, 50 tree-solutions were selected.

4.2.2 DRF Testing

In the testing phase, the accuracy of the trained models was verified using the testing set. Table 4 reports the RMSE and R^2 performance metrics for each variable with additional comparisons between random guessing and specifically guessing the mean values for each of the dependent variables. This part of the experiment enabled an assessment of the model's superiority over various types of educated guesses.

Table 4: RMSE score comparison for each variable between Machine Learning model (ML), Mean Guessing (MG), and random guessing (RG).

Dependent Variable	ML	MG	RG
Achievement	0.1282	0.1376	0.1393
Benevolence	0.0974	0.1046	0.1485
Conformity	0.1267	0.1328	0.1454
Hedonism	0.1056	0.1133	0.1134
Power	0.1195	0.1293	0.1293
Security	0.1134	0.1195	0.1515
Self-Direction	0.1146	0.1180	0.1303
Stimulation	0.1144	0.1182	0.1244
Tradition	0.1031	0.1100	0.1445
Universalism	0.1017	0.1081	0.1086

Furthermore, Figure 3 reports the mean importance of the features across all of the ten basic human values based on the average contribution of each feature to the overall explained variance. Since these scores represent the average contributions across all of the values, it should be noted that certain values can be predicted with higher and lower accuracy, and the cost of obtaining certain demographic features should be considered during data collection.

4.3 Comparison of Approaches

Finally, a comparison between the predictive performance of the two approaches is presented in Table 5, across all of the dependent variables in terms of both the R^2 and RMSE scores. Since the interpretation of R^2 scores is relatively straightforward as the percentage of variability explained in the dependent variable by the independent variables, for the purpose of comparison, this measure of goodness of fit is used. In the case of both approaches, the predictability of Power is the highest, implying that Power can

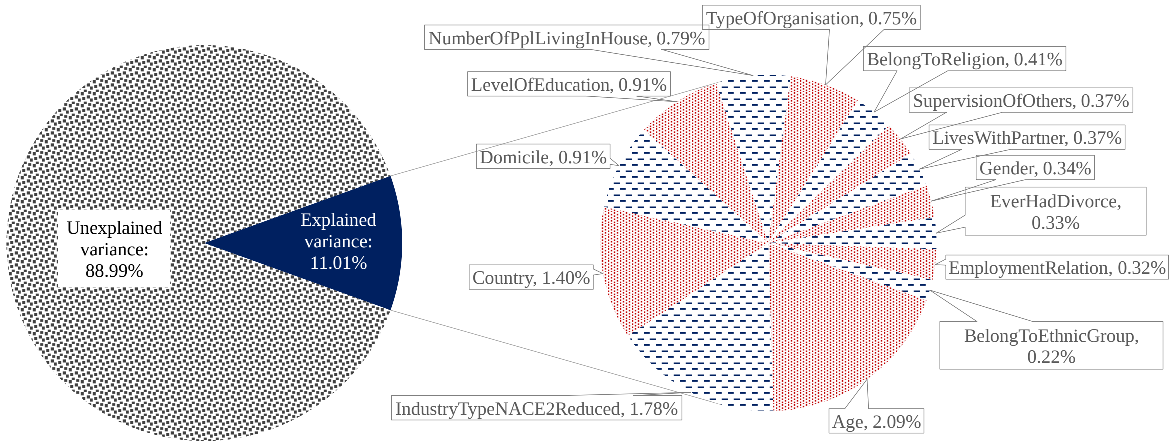

Figure 3: Mean feature importance for predicting the 10 basic human values from observable features by ML approach.

be predicted with the highest accuracy from the available set of demographic variables. On the other hand, Self-direction and Stimulation values are at the lowest end of predictability, which indicates that demographic features are less useful for inferring these particular values. While the LR approach shows slightly better performance than the ML approach in terms of R^2 scores across all of the dependent variables, both data-analytic approaches converge on similar overall results in terms of predictive performance, which further consolidates the findings.

Table 5: Predictive performance comparison of machine learning (ML) and linear regression (LR) approaches in terms of R^2 and RMSE scores.

Dependent Variable	ML approach R^2	ML approach RMSE	LR approach R^2	LR approach RMSE
Achievement	0.13	0.128	0.16	0.127
Benevolence	0.14	0.097	0.16	0.095
Conformity	0.09	0.127	0.11	0.126
Hedonism	0.12	0.106	0.18	0.104
Power	0.15	0.120	0.18	0.118
Security	0.08	0.113	0.12	0.113
Self-Direction	0.07	0.115	0.09	0.113
Stimulation	0.08	0.114	0.09	0.114
Tradition	0.12	0.103	0.14	0.102
Universalism	0.11	0.102	0.13	0.101

5 DISCUSSION

The main objective of this study was to assess the utility of demographic features in predicting stakeholder motivation, operationalized as the basic human values. We have shown through a set of experiments how these observable attributes can be utilized for predicting a subject's motivational profile. The results suggest that the overall predictability of these psychological variables from demographic features is relatively low, but that the usefulness of such assessments is highly dependent on the context in which the results are to be used. In cases where no prior information is available, even a slight reduction in uncertainty can be significant and worth the effort of gathering additional, easily observable features.

A study by Kosinski, Stillwell, and Graepelsing (Kosinski et al., 2013) has demonstrated how a set of psychological constructs (the Big 5 traits) can be predicted from online behavioral traces. Firstly, the study showed that certain differences can be expected among the Big 5 traits in their level of predictability: Openness (r = 0.43), Extraversion (r = 0.40), Neuroticism and Agreeableness (r = 0.3), and Conscientiousness (r = 0.29), covering a range between 8.41 and 18.49 in terms of the R^2. Considering that the present study only relied on demographic features, the level of predictability matched closely, even though behavioral features might convey a lot more information about latent traits. Furthermore, the aforementioned study compared the predictive accuracy obtainable from observable features, to the predictive accuracy achievable by administering the same psychometric instrument for the same respondent at two points in time. The correlation between these scores (test-retest reliability) varies between r = 0.55-0.75, indicating a possible upper bound in terms of the predictability of relatively stable psychological traits by standard, validated instruments.

The experiments conducted with the ML approach established that the model's performance is superior to random guessing, as well as educated guessing (e.g.

a guess of the group means), and that the LR approach had a higher level of performance when using different combinations of predictor variables, but also that most of these differences are only marginal. The differences could be attributed to the automated data preparation in the case of the LR approach, which shows the implementation's additional usefulness during the analysis of complex survey data.

In sum, country, age, and type of industry one is working for are the most important features that can be easily obtained and used for the prediction of the majority of basic values from the available set of features included in the ESS dataset. Therefore, identification and inclusion of other demographic features (which might be more difficult to obtain) do not necessarily provide additional predictive utility. This is important knowledge for an analyst when considering the cost-benefit of gathering a greater amount of descriptive data with the intention of achieving higher accuracy. In order to identify potentially more useful predictor variables, further studies will focus on features that reflect previous choices in a subject's history.

5.1 Legal and Ethical Considerations

It should be noted that there are important legal and ethical aspects when human subjects are involved both in research and in the real-world application of the described profiling method. For this reason it is necessary to outline and separate the conditions under which the method's application can be considered ethical or legal. While the distinction between law and ethics is often unclear, they are fundamentally different (Hvinden et al., 2016). Both are normative, but ethical norms are formulated as guidelines rather than as prescriptions and prohibitions. Ethics is a collection of fundamental concepts and guidelines that informs individuals about desirable actions in certain situations. Legislation, on the other hand, refers to a systematic body of rules and regulations in written form that aim to govern the behavior of individuals within the boundaries of a particular organization (e.g. country) and unlawful activities are penalized and sanctioned. The difference between ethics and law is also expressed in the corresponding documents.

Ethical guidelines (e.g. the Guidelines for Research Ethics in the Social Sciences, Humanities, Law, and Theology (Hvinden et al., 2016)) developed for conducting research with human participants require: respect for human dignity, privacy, safeguarding against harm, compliance with the duty to inform, and the obtaining of the participant's consent, especially in cases where sensitive personal data is collected. There are also exceptions from the main rule concerning informed consent e.g. observation in public arenas, public figures, if the research does not involve direct contact with the participants, and in cases where information cannot be provided before the research is initiated because it would affect the outcomes of the experiment. These exceptions must be justified by proving the they add value to the research and by demonstrating the lack of alternative options.

Laws vary with time and across territories; therefore, it is crucial to have an up-to-date and contextual understanding of the legal regulations concerning any activity. Different laws have been developed for the collection and protection of personal data across nations. Member states of the European Union (EU) and the European Economic Area (EEA) have opted for an all-encompassing regulation named the European General Data Protection Regulation (GDPR) (European Union, 2016). The GDPR requires that the processing of personal (linkable to a person) and sensitive data (health, race or ethnic background, sexuality, political, or religious beliefs) should be done with free and informed consent, and that data processors are required to protect the privacy of respondents, and, therefore, ensure confidentiality. A different approach is used by the United States, which implements various sector-specific data protection laws that work together with state level legislation (e.g. HIPAA, NIST 800-171, the Gramm-Leach-Bliley Act, the Federal Information Security Management Act) (Coos, 2018).

The overview on the legal and ethical aspects aimed to highlight some important issues that have to be taken into consideration when it comes to either the development or the application of any profiling method.

6 CONCLUSIONS

This study aimed at increasing the real-world applicability of the CIRA method that addresses human-related risks within the domain of information security. The method focuses on stakeholder motivation and requires the inference of motivational profiles without direct involvement of the stakeholders. Therefore, we investigated the usefulness of easily observable demographic features for inferring stakeholder motivational profiles. By analyzing a high-quality dataset from representative European samples, and utilizing various data-analytic approaches, we showed that demographic features have some limited usefulness in terms of deriving stakeholder motivation. While the analysis was limited to respondents from European countries, cultural differences account for the majority of variances explained. In sum, these results are useful for characterizing individuals' motivational profiles especially, when limited access to subjects is assumed, and in cases where subjects might be motivated to answer dishonestly to direct questions. While the primary application of these results is the CIRA method of risk analysis, other domains could benefit from predicting inaccessible subject's motivational profiles, especially where decisions are characterized by trade-offs between various objectives and have great potential impact (e.g. intelligence analysis, operations research, etc.). Future work may expand the analysis to include other regions of the world (e.g. USA, Eastern-cultures) to investigate whether the predictability of value profiles is affected by deeper cultural differences. Finally, these findings provide a solid benchmarking baseline for other future work, which will investigate other classes of observable features for inferring motivational profiles. More specifically, observables that represent the outcome of a conscious decision process (e.g. ownership of items, style, etc.) will be analyzed in terms of their capability to provide insight into the decision-maker's value structure.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers for their constructive comments and suggestions for improving the paper.

This work was partially supported by the project IoTSec – Security in IoT for Smart Grids, with number 248113/O70 part of the IKTPLUSS program funded by the Norwegian Research Council.

REFERENCES

Anderson, R. and Moore, T. (2009). Information security: where computer science, economics and psychology meet. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 367(1898):2717–2727.

Arora, J. (2017). Corporate governance: a farce at volkswagen? *The CASE Journal*, 13(6):685–703.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Coos, A. (2018). Eu vs us: How do their data protection regulations square off?

Dembe, A. E. and Boden, L. I. (2000). Moral hazard: a question of morality? *New Solutions: A Journal of Environmental and Occupational Health Policy*, 10(3):257–279.

European Union (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, L119:1–88.

Eurostat, N. (2008). Rev. 2–statistical classification of economic activities in the european community. *Office for Official Publications of the European Communities, Luxemburg*.

Forbes, D. L. (2011). Toward a unified model of human motivation. *Review of general psychology*, 15(2):85.

Franklin, Z. (2014). Justice for revenge porn victims: Legal theories to overcome claims of civil immunity by operators of revenge porn websites. *California Law Review*, pages 1303–1335.

Goldberg, L. R., Sweeney, D., Merenda, P. F., and Hughes Jr, J. E. (1998). Demographic variables and personality: The effects of gender, age, education, and ethnic/racial status on self-descriptions of personality attributes. *Personality and Individual differences*, 24(3):393–403.

Grace-Martin, K. (2008). Assessing the fit of regression models. [Online; accessed 05-July-2018].

Greitzer, F. L. and Hohimer, R. E. (2011). Modeling human behavior to anticipate insider attacks. *Journal of Strategic Security*, 4(2):25–48.

Gudaitis, T. M. (1998). The missing link in information security: Three dimensional profiling. *CyberPsychology & Behavior*, 1(4):321–340.

H2O.ai (2018a). *Distributed Random Forest (DRF)*.

H2O.ai (2018b). *H2O.ai*.

Hvinden, B., Johanne Bang, K., Fjørtoft, K., Holand, I., Johnsen, R., Kolstad, I., Monsen, T., Nevøy, A., Sandmo, E., Skilbrei, M.-L., Staksrud, E., Tande, K. M., Ulleberg, P., Øyum, L., and Enebakk, V. (2016). Guidelines for research ethics in the social sciences, humanities, law and theology.

Knafo, A. and Sagiv, L. (2004). Values and work environment: Mapping 32 occupations. *European Journal of Psychology of Education*, 19(3):255–273.

Kosinski, M., Stillwell, D., and Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, page 201218772.

Krebs, B. (2012). Fbi: Smart meter hacks likely to spread. [Online; accessed 26-June-2018].

Kulik, B. W., O'Fallon, M. J., and Salimath, M. S. (2008). Do competitive environments lead to the rise and spread of unethical behavior? parallels from enron. *Journal of Business Ethics*, 83(4):703–723.

Liebowitz, S. J. and Margolis, S. E. (1994). Network externality: An uncommon tragedy. *Journal of economic perspectives*, 8(2):133–150.

N.A. (2018). European social survey round 8 data. *NSD - Norwegian Centre for Research Data, Norway - Data Archive and distributor of ESS data for ESS ERIC*.

Nagelkerke, N. J. et al. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 78(3):691–692.

Patro, S. and Sahu, K. K. (2015). Normalization: A preprocessing stage. *arXiv preprint arXiv:1503.06462*.

Rajbhandari, L. and Snekkenes, E. (2013). Using the conflicting incentives risk analysis method. In *IFIP International Information Security Conference*, pages 315–329. Springer.

Schwartz, S. (2016). Computing scores for the 10 human values.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In *Advances in experimental social psychology*, volume 25, pages 1–65. Elsevier.

Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values? *Journal of social issues*, 50(4):19–45.

Schwartz, S. H. (2007). Basic human values: theory, methods, and application. *Risorsa Uomo*.

Schwartz, S. H. (2012). An overview of the schwartz theory of basic values. *Online readings in Psychology and Culture*, 2(1):11.

Schwartz, S. H. (2013). Culture matters: National value cultures, sources, and consequences. In *Understanding Culture*, pages 137–160. Psychology Press.

Schwartz, S. H. and Rubel, T. (2005). Sex differences in value priorities: Cross-cultural and multimethod studies. *Journal of personality and social psychology*, 89(6):1010.

Snekkenes, E. (2013). Position paper: Privacy risk analysis is about understanding conflicting incentives. In *IFIP Working Conference on Policies and Research in Identity Management*, pages 100–103. Springer.

Szekeres, A. and Snekkenes, E. A. (2018). Unobtrusive psychological profiling for risk analysis. In *Proceedings of the 15th International Joint Conference on e-Business and Telecommunications - Volume 1: SECRYPT,*, pages 210–220. INSTICC, SciTePress.

Yang, H. (2013). The case for being automatic: introducing the automatic linear modeling (linear) procedure in spss statistics. *Multiple Linear Regression Viewpoints*, 39(2):27–37.

APPENDIX

	df regression	*df* residual	*F*	*adjusted* R^2	*Intercept*	*Unstandardized Beta*	*Country (coded as)*	*Age*		*Gender (coded as)*
Achievement	11	22,036	390.72	0.16	0.52	0.08	Finland (0)			
						0.15	United Kingdom (1)			
						0.15	Lithuania (2)			
						0.14	Netherlands (3)			
						0.01	Sweden (4)			
						0.08	Belgium, Switzerland (5)	-0.002		
						0.06	Spain, Poland (6)			
						0.10	Austria, Estonia, Italy, Russian Federation (7)			
						0.11	Czech Republic, Ireland (8)			
						0[a]	Iceland, Norway (9)			
						0.05	Hungary, Slovenia (10)			
Benevolence	13	22,034	329.04	0.16	0.64	-0.10	Finland (0)		-0.03	Male (1)
						-0.04	United Kingdom (1)		0[a]	Female (2)
						-0.06	Lithuania (2)			
						0.03	Netherlands (3)			
						-0.12	Sweden (4)			
						-0.10	Belgium, Switzerland (5)			
						-0.05	Spain, Poland (6)			
						-0.03	Austria, Estonia, Italy, Russian Federation (7)			
						-0.02	Czech Republic, Ireland (8)			
						0[a]	Iceland, Norway (9)			
						0.01	Hungary, Slovenia (10)			
						-0.08	Germany, France, Israel, Portugal (11)			
Conformity	9	22,038	303.35	0.11	0.51	-0.10	Finland (0)			
						-0.03	United Kingdom (1)			
						-0.07	Lithuania (2)			
						-0.05	Netherlands (3)			
						-0.09	Sweden (4)	0.002		
						-0.06	Belgium, Switzerland (5)			
						-0.08	Spain, Poland (6)			
						-0.11	Austria, Estonia, Italy, Russian Federation (7)			
						0[a]	Czech Republic, Ireland (8)			
Hedonism	13	22,034	341.90	0.18	0.60	-0.01	Finland (0)			
						-0.11	United Kingdom (1)			
						-0.07	Lithuania (2)			
						-0.13	Netherlands (3)			
						-0.05	Sweden (4)			
						0.01	Belgium, Switzerland (5)	-0.002		
						-0.06	Spain, Poland (6)			
						0.00	Austria, Estonia, Italy, Russian Federation (7)			
						-0.03	Czech Republic, Ireland (8)			
						0[a]	Iceland, Norway (9)			
						-0.04	Hungary, Slovenia (10)			
						0.02	Germany, France, Israel, Portugal (11)			
Power	9	22,038	510.15	0.18	0.46	0.09	Finland (0)			
						0.13	United Kingdom (1)			
						0.05	Lithuania (2)			
						-0.03	Netherlands (3)			
						0.12	Sweden (4)			
						0.02	Belgium, Switzerland (5)			
						0.06	Spain, Poland (6)			
						-0.05	Austria, Estonia, Italy, Russian Federation (7)			
						-0.02	Czech Republic, Ireland (8)			
						0[a]	Iceland, Norway (9)			

Note. [a] *reference variable; all SE B < .005; for all included variables p < .05*

Figure 4: Final regression models for each dependent variable (1/2).

	df					Unstandardized Beta										
	regression	residual	F	adjusted R^2	Intercept		Country (coded as)	Age		Religion (coded as)	Level of education		NACE classification of economic activities (coded as)		Employment Relation (coded as)
Security	12	272,374	267.72	0.12	0.51	0.03	Finland (0)								
						0.10	United Kingdom (1)								
						-0.02	Lithuania (2)								
						-0.04	Netherlands (3)								
						0.08	Sweden (4)								
						0.08	Belgium, Switzerland (5)	0.001			-0.01				
						0.02	Spain, Poland (6)								
						0.04	Austria, Estonia, Italy, Russian Federation (7)								
						0.05	Czech Republic, Ireland (8)								
						0^a	Iceland, Norway (9)								
						0.07	Hungary, Slovenia (10)								
Self-Direction	11	22,036	212.09	0.09	0.47	0.01	Finland (0)							0.04	Self-employed (0)
						-0.01	United Kingdom (1)							0^a	Employee, Working for own family business (1)
						0.03	Lithuania (2)								
						0.06	Netherlands (3)								
						0.06	Sweden (4)				0.01				
						-0.04	Belgium, Switzerland (5)								
						0.04	Spain, Poland (6)								
						0.04	Austria, Estonia, Italy, Russian Federation (7)								
						0.01	Czech Republic, Ireland (8)								
						0^a	Iceland, Norway (9)								
Stimulation	6	22,041	361.72	0.09	0.56	-0.02	Finland (0)								
						-0.04	United Kingdom (1)								
						-0.03	Lithuania (2)	-0.002							
						-0.01	Netherlands (3)								
						-0.04	Sweden (4)								
						0^a	Belgium, Switzerland (5)								
Tradition	9	22,038	415.14	0.14	0.54	-0.03	Finland (0)		0.04	Yes (1)					
						-0.05	United Kingdom (1)		0^a	No (2)					
						-0.05	Lithuania (2)								
						-0.01	Netherlands (3)	0.001			-0.01				
						-0.02	Sweden (4)								
						-0.03	Belgium, Switzerland (5)								
						0^a	Spain, Poland (6)								
Universalism	16	22,031	209.70	0.13	0.51	-0.02	Finland (0)					0.01	Mining and quarrying (0)		
						-0.09	United Kingdom (1)					-0.01	Water supply; sewerage, waste management and remediation activities, Information and communication (2) Professional, scientific and technical activities, Administrative and support service activities (2)		
						-0.08	Lithuania (2)								
						-0.04	Netherlands (3)					0.02	Agriculture, forestry and fishing, Manufacturing, Accommodation and food service activities (3)		
						-0.05	Sweden (4)					0.03	Public administration and defence; compulsory social security, Education (4) Other service activities, Activities of extraterritorial organisations and bodies (4)		
						-0.04	Belgium, Switzerland (5)	0.001							
						-0.10	Spain, Poland (6)					-0.004	Human health and social work activities, Arts, entertainment and recreation, Activities of households as employers (5)		
						-0.01	Austria, Estonia, Italy, Russian Federation (7)					0^a	Electricity, gas, steam and air conditioning supply, Wholesale and retail trade; repair of motor vehicles and motorcycles (6) Transportation and storage, Financial and insurance activities, Real estate activities (6)		
						-0.04	Czech Republic, Ireland (8)								
						0^a	Iceland, Norway (9)								
						-0.07	Hungary, Slovenia (10)								

Note. [a] *reference variable; all SE B < .005; for all included variables p < .05*

Figure 5: Final regression models for each dependent variable (2/2).