

Criminal Trend Analysis of Illegal Use of Personal Data for a Period of 2015-2019

Nikola Protrka*, Mihael Plečas**

* Police College, Zagreb, Croatia

** Libertas International University, Zagreb, Croatia

nprotrka@fkz.hr, mplecas@libertas.hr

Abstract - This paper presents an analysis of the criminal offence of the unauthorised use of personal data in the Criminal code of the Republic of Croatia, described in article 146 (including the incidence of these crimes over a five-year period).

The paper analyses the impact and efficiency of criminal offences under the code for a five-year period from 2015-2019 by reviewing data related to reported criminal offences with detailed statistics about reported, resolved and subsequently discovered offences of illegal use of personal data in the Republic of Croatia. Authors have chosen this specific period of time due to the fact that the European General Data Protection Regulative (GDPR) became active in 2018 and it is directly obliging all the member states to pursue consistent implementation regarding personal data. *Inter alia*, an analysis was performed on victims as well as perpetrators of relevant criminal offences and shown also by the gender. The objectives of the research were to discover how efficient were Croatian institutions in solving criminal offences that are related to illegal use of personal data in the Republic of Croatia 2015-2019. The paper also describes descriptive statistical data concerning these offences and give us prediction of a number of possible offences for a following year.

Keywords - personal data; analysis; trend; illegal use of personal data; criminality; GDPR

I. INTRODUCTION

The supervisory authority for processing and protection of personal data in Croatia is the Croatian Personal Data Protection Agency, whilst criminal prosecution is carried out by the police and State Prosecution Office. On a daily base we are exposed to possible illegal use of our personal data. We store personal data on different storage media, including cloud storage and we use it on a daily basis in wide spectrum of services from social media, online shops, filling in different kinds of forms for credit cards, loyalty or discounts etc. Personal devices have had a significant increase in storage over the years, and there can be stored a lot of our personal data. Same situation is going on with cloud storage [1].

Internet growth statistics show that number of individuals who are online has increased from 2000 to 2018 by 10520% [2]. It is very clear that during that period of time the use of Internet connected devices has grown rapidly, and cybercrime has become part of everyday life [3].

In Croatian legislation (Criminal Code, 2011) there is a specific article covering the illegal use of personal data, described in article 146 named Illegal Use of Personal Data. Paragraph 1 describing that whoever, contrary to the conditions set out by the law, collects, processes or uses personal data of physical persons can be punished by imprisonment, while in next paragraph of the same article is described that whoever, contrary to the conditions set out by the law, transfers personal data outside of the Republic of Croatia for further processing, or makes them public or in some other way available to a third party, or whoever by the act referred to previous paragraph acquires considerable material gain for himself or herself or for another or causes considerable damage can be punished by imprisonment [4]. By enacting the European General Data Protection Directive there has been a significant influence on the perception of this criminal offence through different activities not only by Croatian Personal Data Protection Agency but other state and private agencies throughout different educations, seminars, guidelines and of course by legal regulations [5,6].

This paper analyses statistics of Ministry of Interior of the Republic of Croatia of these legal provisions for period from 2015 to 2019., and it's related to illegal use of personal data and the number of reported criminal offences under article 146 of the Criminal Code [7,8,9,10,11].

In conducting forensic investigations and trying to find out what have happened with illegal use of personal data, a number of factors could change the results of report. Usually, we are dealing with digital evidence or artefacts [12]. For that misuse of data, we most commonly find out when personal data was misused such as identity theft, selling for direct marketing or simply to wait for the best opportunity to for implementing of some sort of illegal activity, in other words, criminalisation of personal data [13].

II. ILLEGAL USE OF PERSONAL DATA

Due to developing technology and development of computer networks that are connecting computers and other devices that we use on an everyday basis, and are colloquial known as Internet of Things (IoT – can be described like the network of physical hardware objects or “things” that are embedded with sensors, built-in

software, for the purpose of connecting and exchanging data with other devices and systems over the computer network), regularly we are facing an unimaginable amount of personal data. As the numbers of such devices grow so do the risks of using this technology. On a daily base we need to raise an awareness of protecting data that these devices gather and transmit for further processing [14].

The European General Data Protection Regulation defines data in Art. 4 and states that this is any information relating to an identified or identifiable natural person (data subject); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier or to one or more factors specific to the physical, psychological, genetic, mental, economic, cultural or social identity of that natural person [15, 16].

The development of databases that include personal data has been very fragmented, without too much thinking or planning about mutual compatibility [17]. Such interconnectedness of the system was noticed by the European Union and in 2019 two regulations were passed [18].

In the Republic of Croatia, the criminal offense described in Act. 146 of the Criminal Code is most often mentioned when it comes to the unauthorized or illegal use of personal data. This offense is a one of other offenses that belongs to criminal offenses against privacy in the Criminal Code. Here, we must mention other criminal offenses against privacy, namely: Article 141 - Violation of the inviolability of Home and Business Premises, Article 142 - Violation of the Privacy of Correspondence and Other Parcels, Article 143 - Unauthorised Audio Recording and Eavesdropping, Article 144 - Unauthorised Taking of Pictures, Article 145 - Unauthorised Disclosure of a Professional Secret [4].

The Republic of Croatia has (as a consequence) taken steps to effectively undertake counter-action activities against the illegal use of personal data [19]. The Ministry of Interior has presented the results of work on related problems in statistical representations that were used for further analysis. Table I. shows the total number of reported and solved cases of criminal offences of illegal use of personal data in the Republic of Croatia over the last five-year period. Likewise, it shows the number of subsequently discovered criminal offences from previous years.

TABLE I. REPORTED, RESOLVED AND SUBSEQUENTLY DISCOVERED OFFENCES OF ILLEGAL USE OF PERSONAL DATA IN THE REPUBLIC OF CROATIA

Year	Reported criminal offences	Resolved criminal offences	Subsequently discovered criminal offences
2015.	601	600	580
2016.	482	476	464
2017.	270	260	242
2018.	228	196	175
2019.	289	269	188

Reference/source: Processed by authors based on the data of the Ministry of the Interior

Based on data in table II. shows the overall number of victims, and by the gender of the harmed individual or victim. There is a visible difference in number of male victims to in a way that male victims make up majority.

TABLE II. HARMED PERSONS (VICTIMS/TARGETS)

Year	Total of harmed individual	Natural persons	Female	Male
2015.	623	609	168	441
2016.	460	455	153	302
2017.	272	229	74	155
2018.	230	203	75	128
2019.	308	306	109	197

Reference/source: Processed by authors based on the data of the Ministry of the Interior

In table III. the number of perpetrators of criminal offences is shown and systematically analysed by type and gender. It is evident that in the observed period there is a larger number of male than female perpetrators of these criminal offences.

TABLE III. PERPETRATORS OF CRIMINAL OFFENCE OF ILLEGAL USE OF PERSONAL DATA

Year	Total of perpetrator	Natural persons	Female	Male
2015.	124	124	31	93
2016.	108	108	27	81
2017.	83	83	22	70
2018.	70	70	19	64
2019.	74	74	16	54

Reference/source: Processed by authors based on the data of the Ministry of the Interior

III. ANALYSIS AND INTERPRETATION OF THE RESULTS OF THE RESEARCH

A. Description of the researched data

Data on the total number of reported offences of illegal use of personal data in the Republic of Croatia in period from 2015-19 are shown in Table IV. In this period the number of cases has decreased from 601 in 2015 to 289 in 2019. Observed values, the number of reported offences decreased each year according to the average growth rate, i.e., the geometric mean, -16,7267 cases. The arithmetic mean of the growth rate is somewhat lower and is -13,1462, which is unusual as the geometric mean is generally less than the arithmetic mean.

TABLE IV. CRIMINAL OFFENCES OF ILLEGAL USE OF PERSONAL DATA IN THE REPUBLIC OF CROATIA

Year	Number of criminal offences	Chain index	Growth rate %
2015.	601		
2016.	482	80,1997	-19,8003
2017.	270	56,0166	-43,9834
2018.	228	84,4444	-15,5556
2019.	289	126,7544	26,7544
Arithmetic mean of the growth			-13,1462
Average growth rate (geometric mean)			-16,7267

Reference/source: Processed by authors based on the data of the Ministry of the Interior

In Table IV we can see the values of selected indicators of descriptive statistics. Variable values *Number of observations* is 5. This relates to five years in which the numbers of reported criminal offences related to illegal use of personal data were observed.

TABLE V. INDICATOR VALUES OF DESCRIPTIVE STATISTICS.

Indicator of descriptive statistics	Value
Number of observations	5
Minimum	228
Median	289
Maximum	601
Range of variation	373
Arithmetic mean	374
Standard deviation	160,1171
Coefficient of variation	42,81 %

Reference/source: Processed by authors based on the data of the Ministry of the Interior

Variable values of *Minimum* amounts to 228 and it is related to number of criminal offences in year 2018., apropos the penultimate year of the observed period. Variable values *Median* amounts 289 and it is related to number of criminal offences in year 2019 (the last year of observed period.)

Variable values *Maximum* amounts 601 and it is related to the number of criminal offences in year 2015. apropos the first year of the data.

Variable values *Range of variation* amounts 373. The arithmetic mean of reported criminal offences that relates to illegal use of personal data is 374 cases. Variable values *Standard deviation* that relates to average aberration from the arithmetic mean is 160,1171. Based on the value of the *Coefficient of variation*, which refers to variability of the observed phenomenon amounts 42,81%, a conclusion is drawn that variability is moderate.

Taking into account the decline in numbers of reported offences that relate to illegal use of personal data in first four years of observed period, it is justified to develop trend model to predict numbers in the coming years. Since it is that in the period of 2015 – 2019, variability of the trend in reporting criminal offences was moderate, which confirms variable values of *Coefficient of variation* in table 5. that amounts 42,81%, it is legitimate to create two trend models. One should apply linear regression analysis, and the other, an exponential regression analysis. Both of these models are based on the calculation of least squares difference in relation to actual values [20].

B. Description of trend model of number movements in reported criminal offences

Results of linear regression analysis of number occurrence of reported criminal offences that relate to illegal use of personal data are presented in table VI. For a description of the trend model, in table VI., the most important variables are *P value*, *Intercept*, *Year* and *Adjusted R²*.

Whether or not the calculated values of the linear regression analysis should be calculated, that is, the linear trend model considered statistically significant depends on the value of variable *P value*, which refers to the probability of wrongly derived conclusion in regression analysis [21].

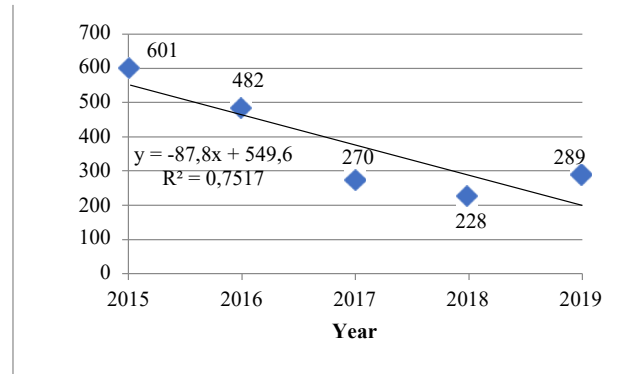
Because the variable value is 0,0570, which is slightly higher than the value limit of 0,05 for a less rigid test of statistical value, all the values presented in the table cannot be considered of having statistical significance. The same conclusion can be drawn based on variable value Adjusted R^2 which refers to corrected coefficient of determination that amounts 0,6690 meaning that 66,90% of occurrence is described with linear trend model of number movements shown that are related to reported criminal offences of illegal use of data protection. Considering that linear regression analysis was conducted on a small sample (5), conclusions of the representativeness of the trend model are made based on variable value of R^2 which refers to coefficient of determination which determines percentage of described occurrence when observed sample is significant number [22].

TABLE VI. LINEAR REGRESSION ANALYSIS

Regression Statistics			
Multiple R	0,8670	Pearson's coefficient of linear correlation	
R^2	0,7517	Coefficient of determination	
Adjusted R^2	0,6690	Corrected coefficient of determination	
Standard Error	92,1260	Standard deviation of regression	
Observations	5		
ANOVA			
	<i>df</i>	<i>F</i>	<i>Significance F</i>
Regression	1	9,0829	0,0570
	<i>Coefficients</i>	<i>andardError</i>	<i>P-value</i>
Intercept	549,6	71,3605	0,0045
Year	-87,8	29,1328	0,0570

Reference/source: Processed by authors based on the data of the Ministry of the Interior

Variable value *Intercept* relates to value of constant and variable value *Year* relates to value of regression coefficient. Based on these two values from table VI., regression equation of trend model was conducted and is: $y = 549,6 - 87,8 \cdot X$.



FIG' I. LINEAR TREND MODEL OF A DECREASE IN INCIDENTS OF CRIMINAL OFFENCES RELATED TO ILLEGAL USE OF PERSONAL DATA

Reference/source: Processed by authors based on the data of the Ministry of the Interior

The value of the constant is 549,6 and presents a start value of the number of criminal offences in the linear trend model of the observed occurrence. In as much as the real value in the first year of observed period is 601, deviation of the start value in trend model is 9,35 % which is considered unacceptable. The value of the regression coefficient is -87,8. Since it's negative, the number of reported offences of illegal use of personal data per year is on average reducing to 88.

The linear trend model of number movement of reported criminal offences related to illegal use of personal data has been shown graphically in fig. I. Regression equation on the graph has identical values of constant and regression coefficient to the ones from table VI. Also presented is the value of the coefficient of determination or variable R^2 that amounts 0,7517 which indicates how representative the trend model is. However, due to only 5 observed values, for observed regression analysis it is relevant variable value of *Adjusted R^2* that is the value of the corrected coefficient of determination from table VI. Based on the regression equation that is indicated in the figure, in 2020 the expected number of criminal offences is 110. The authors consider this value to be an insufficiently good prediction. In 2020 can be expected more than 200 cases due to huge amount of employed people worked from home due to the COVID-19 pandemic. In 2021, the expected number of criminal offenses is 23, which the authors consider an unusable prediction of the analysed occurrence.

By conducting the linear regression it has been determined that the calculated values cannot be considered statistically significant because the value of the variable *P-value* in Table IV, that refers to probability of erroneously derived conclusion on significance is 0,0570 meaning that is slightly higher than borders value for less rigorous test of statistical significance, what confirms the value of the variable *Adjusted R^2* in table IV. which refers to the corrected coefficient of determination.

This is the percentage by which has the linear trend model successfully described observed occurrence is 66,90% which is less than acceptable percentage, so it is

justifiable to describe the observed occurrence by another regression model [23].

The exponential trend model of the occurrence of the numbers of the reported criminal offences related to illegal use of personal data is graphically presented in Figure 2. It is also presented the value of the variable R^2 variable *Adjusted R^2* , which refers to a corrected coefficient of determination of 0,7517 and which indicates how representative the trend model is. However, due to only 5 observed values for the regression analysis, i.e. trend model, the relevant is the value of the variable is *Corrected coefficient of determination*. Based on the values marked on the graph of the constant and regression coefficient in 2020 the expected number of reported criminal offences related to illegal use of personal data is 110.

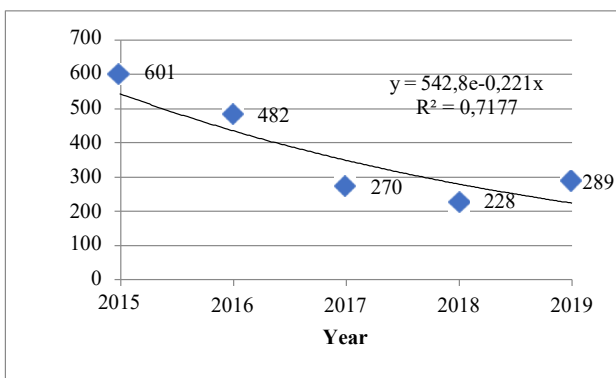


FIG. II. EXPONENTIAL TREND MODEL OF REDUCTION OF INCIDENCE OF OFFENCES RELATED TO ILLEGAL USE OF PERSONAL DATA

Reference/source: Processed by authors based on the data of the Ministry of the Interior

Figure II. presents regression equations of the exponential trend model of the occurrence of criminal offences of illegal use of personal data $Y = 542,8 \cdot e^{0,221x}$.

The value of the constant is 542,8 and it represents the initial trend value of the number of criminal offences in exponential trend model of the observed occurrence. As the actual value in the first year of the observed period is 601, the deviation of the initial value in trend model is 11,27 % which is considered unacceptable. The value of the regression coefficient is - 0,221. Because it is negative, it indicates that the number of reported criminal offences related to illegal use of personal data is decreasing yearly in average for 22,1%.

The value of the regression coefficient from the exponential trend model should be similar to the value of the geometric mean of chain index which is described the movement of the numbers of reported cases related to the misuse of personal data. By comparing the value of the geometric mean of the chain index from Table 4, from which the average annual rate of decline in the number of reported cases is -16,73 % with the value of the regression coefficient from fig. 2. is 22,10 %. Consequently, there is a significant deviation in absolute values of 5,37 %, or relatively more than 30 %.

The value of the coefficient of determination R^2 which is 0,7177 indicates how representative the trend model is. The value indicates that 71,77 % of the occurrence of the number of criminal offences related to illegal use of personal data is described by the exponential trend model. Since this value is less than the value of the corrected coefficient of determination of the linear trend model, the exponential trend model is less representative for the description of the observed phenomenon than the linear trend model. This value of the constant and regression coefficient, the calculated number of reported offences related to illegal use of personal data in 2020 is 180. As the authors mentioned before, in 2020 can be expected more than 200 cases due to huge amount of employed people worked from home due to the COVID-19 pandemic. So the predicted value the authors consider an unusable prediction. The same applies for the prediction in 2021 because the expected number of criminal offenses is 144. Both prediction models are not reliable enough because after four years of decreasing of the cases of criminal offences of illegal use of personal data, there has been an increase of cases in the fifth year which significantly affected on trend models.

After the paper was written, the data for 2020 have become available. The registered number of the cases of criminal offences of illegal use of personal data in 2020 was 269 [24]. That number is greater than 200 as predicted by the authors and less than number of cases in 2019. It means that number of cases of criminal offences of illegal use of personal data in 2019 represents an exception in trend of decreasing number of the cases. The authors expect that number of the cases in 2021 will be less than number of cases in 2020.

Once the data for 2021 becomes available, the authors plan to develop new trend models to make these models more reliable in predicting the number of cases of criminal offences of illegal use of personal data. Unfortunately, the exception of the number of cases in 2019 will continue to affect the reliability of the case prediction model.

IV. CONCLUSION

By a visible decrease in the number of reported criminal offences of illegal use of personal data in the observed five-year we can conclude that the trend in the number of the reported criminal offences is declining.

As we can see in research from Goudriaan et al., many crimes are not reported to the police, and we can be sure that in research explained in this paper that dark figure of crime can be high, because sometimes victims were not aware that they were victim of committed crime, as we explained in previous text. There are few factors that influence crime reporting, like economic factors, psychological factors, neighborhood factors and in recent times socio-demographics factors [25]. If we have clearly defined procedure for reporting this type of crime, and citizens are aware of influence and possibilities of illegal use of personal data, then we can hope that the dark figure of this type of crime in near future can be negligible [26].

This can be explained by raising citizens awareness of this issue. The number of solved criminal offences is almost identical to the number of reported ones, and thus we propose that the enforcement activity is in a satisfactory proportion. Certainly, the success of the detection activity should include the subsequently discovered criminal offences.

REFERENCES

- [1] A. Shalaginov (2020) Future Generation Computer Systems: Big Data Analytics and Artificial Intelligence for Cyber Crime Investigation and Prevention. Vol 109. pp 1-2
- [2] H. Jahankhani (2018) Advanced Sciences and Technologies for Security Applications: Cyber Criminology - Crime Data Mining, Threat Analysis and Prediction. Pp 183-202
- [3] M. Tcherni, A. Davies, G. Lopes, A. Lizotte, (2015) The Dark Figure of Online Property Crime: Is Cyberspace Hiding a Crime Wave?. Taylor and Francis Group. Justice Quarterly. pp 2-22
- [4] Criminal code, The Official Gazette of the Republic of Croatia, Narodne novine 125/11, 144/12, 56/15, 61/15, 101/17, 118/18, 126/19
- [5] 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), 2016.
- [6] National GDPR implementation law, Narodne novine 42/18
- [7] Ministry of interior of the Republic of Croatia, Statistical review of safety indicators and the results of work, 2015. Zagreb, 2016.
- [8] Ministry of interior of the Republic of Croatia, Statistical review of safety indicators and the results of work, 2016. Zagreb, 2017.
- [9] Ministry of interior of the Republic of Croatia, Statistical review of safety indicators and the results of work, 2017. Zagreb, 2018.
- [10] Ministry of interior of the Republic of Croatia, Statistical review of safety indicators and the results of work, 2018. Zagreb, 2019.
- [11] Ministry of interior of the Republic of Croatia, Statistical review of safety indicators and the results of work, 2019. Zagreb, 2020.
- [12] O. Yeboah-Boateng, E. Akwa-Bonsu (2016) Digital Forensic Investigations: Issues of Intangibility, Complications and Inconsistencies in Cyber-crimes. In: Journal of Cyber Security and Mobility Vol. 4. Iss 2. pp 87-104
- [13] N. Protrka, (2014) Normative arrangement of personal data protection in the Republic of Croatia. Policija i sigurnost, 22 (4/2013), pp. 509-521.
- [14] G. Vojković, M. Milenković & T. Katulić (2020) IoT & Smart Home Data Breach Risks from the Perspective of Data Protection & Information Security Law. Business systems research, 11 (3), pp. 167-185 doi:10.2478/bsrj-2020-0033.
- [15] G. Vojković (2019) General regulation on protection of personal data. In: T. Velki & K. Šolić (ur.) Challenges of digital world. Osijek, Josip Juraj Strossmayer University of Osijek, Faculty of Education, pp. 177-199.
- [16] J. Čizmić & M. Boban (2018) The effect of the new EU directive 2016/679 (GDPR) on protection of personal data in the Republic of Croatia. Corpus of the Law University in Rijeka, 39 (1), pp. 377-406 doi:10.30925/zpfsr.39.1.13.
- [17] A. Giannakoula, D. Lima, M. Kaiafa-Gbandi. (2020) Combating Crime in the Digital Age: A Critical Review of EU Information Systems in the Area of Freedom, Security and Justice (Challenges for Criminal Law and Personal Data Protection), pp. 27-33
- [18] K. Eisele (2018) Interoperability between EU information systems for security, border and migration management (European Parliamentary Research Service February 2018) PE 615.649.
- [19] D. Dragičević; N. Gumzej. (2015) Privatnost i zaštita osobnih podataka u digitalnom okruženju. In: Dragičević, D. (ur.) Legal Informatics and Information Technologies Law. Zagreb, Narodne novine, pp. 101-164.
- [20] M. Papić, "Applied statistics in MS Excelu", Zagreb, Croatia. Zoro d.o.o., 2014
- [21] B. Petz, (2012) Petz statistics, Jastrebarsko. Croatia. Naklada Slap.
- [22] A. Rozga, B. Grčić, (2009) Business statistic. Split. Croatia. Economy university in Split.
- [23] I. Šošić, V. Serdar, (1995) Introduction to statistics. Zagreb. Croatia. Školska knjiga.
- [24] Ministry of interior of the Republic of Croatia, Statistical review of safety indicators and the results of work, 2020. Zagreb, 2021.
- [25] H. Goudriaan, JP. Lynch and P. Nieuwbeerta (2004) Reporting to the police in western nations: A theoretical analysis of the effects of social context. Justice Quarterly 21(4): 933–969.
- [26] LawTeacher. (2013) The Dark Figure of Crime. [online]. Available from: <https://www.lawteacher.net/free-law-essays/criminology/the-dark-figure-of-crime.php?vref=1> [Accessed 29 March 2021].