

### **Ethical considerations in the current research project**

In the current paper the appropriate care and consideration has been given to the ethical concerns regarding the scraping, processing, and the presentation of the data. Drawing upon the European Commission's (2018) ethical guidelines and principles for ethical conduct in social data science, the potential harm to users of Boliga's web-page were carefully considered. As a step to prevent the mosaic effect and in an effort to anonymize the scraped data only aggregated data will be presented in this paper, so that no single observation can be identified from the analysed data.

No informed consent has been obtained from the users of the site, prompting us to consider the consequences of the lack thereof, as informed consent is paramount to the proper, ethical conduct in social science. However, sometimes, as in this instance, informed consent can be logistically impossible to collect from all participants in the study. Salganik (2017) mentions that informed consent for everything is an ideal, but in practice impossible to obtain and researchers should instead strive to follow an alternative rule, that he describes as: "some form of consent for most things." (p.281) Adhering to this, more complex understanding of the practicality of informed consent, we chose to contact Boliga to inform them of our intent to scrape their website and use the data in an educational context. Boliga responded positively to our inquiry, which we took as informed consent from a third party on behalf of the participating users in our study.

In considering the legal ramifications of our research and to make sure we adhere to the seven principles of GDPR, and other appropriate legislation and legal contracts, we consulted the general guidelines introduced by the Consumer Data Research Center (2018). In particular we noted that we are justified to collect and use the data on the lawful basis of legitimate interests. Furthermore, we consulted with Boliga's Terms and Conditions to avoid any legal ramifications and the appropriate contractual terms of interest can be seen in appendix XX. In order to comply with these terms we refrained from burdening their website's performance by implementing a `time.sleep` function, which causes each scraping iteration to pause for 0.5 seconds before commencing on scraping the next page (see appendix XX). Furthermore, Boliga prohibits the use of automated scrapers and bots, which we did not do, as our scraping was done in a specified time-frame and we did not automate the procedure to be done on multiple occasions.