# ETHICAL GUIDELINES FOR NUDGING IN INFORMATION SECURITY & PRIVACY

## Nudging in Security



There has recently been an upsurge of interest in the deployment of behavioural economics techniques in the information security and privacy domain. One particular intervention, the **nudge**, and the way it exercises its influence is of interest here. We contemplate the ethical ramifications of nudging, in its broadest sense, deriving general principles for ethical nudging from the literature. We extrapolate these principles to the deployment of nudging in information security and privacy.

### DEFINITIONS

### *Nudge:*

Any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives

### *Ethical Principles:*

P1. Respect for Persons

P2. Beneficence

P3. Justice

P4. Scientific Integrity

P5. Social Responsibility

## Why Ethics?

As the nudge becomes the topic of more experiments, and is deployed across public life, its ethical ramifications should be contemplated. Researchers wishing to experiment with the behavioural change efficacy of specific nudges need guidance about how to conduct such experiments in an ethical manner. As Info-S&P researchers ourselves, we focus on the ethics of nudging in Info-S&P.

# P1: Respect

**P1(a). Are all original choices still available? If not, has the withdrawal of some of the options been well-argued?**

P1(b). Will nudgees be aware that an experiment is under way? If not, is the need for this level of deception justified?

P1(c). Will nudgees be aware of the nudge? If not, has the use of a simple nudge been well argued and motivated?

# P2: Beneficence

P2(a). Is the argument for a benefit of applying the nudge well argued, either to the nudgee or to society at large? In particular, have the proposers shown that their nudge is not actually a prod or sludge?

P2(b). Will nudgees be able to contact the choice architect if they have questions or concerns? If yes, is there a commitment to respond to questions within a certain period of time? If not, is there an explanation for this?

P2(c). Has the benefit for the targeted group already been evaluated and reported in the research literature? If not, how will the assumed benefit be evaluated?

# P3: Justice

P2(a). Is the argument for a benefit of applying the nudge well argued, either to the nudgee or to society at large? In particular, have the proposers shown that their nudge is not actually a prod or sludge?

P2(b). Will nudgees be able to contact the choice architect if they have questions or concerns? If yes, is there a commitment to respond to questions within a certain period of time? If not, is there an explanation for this?

P2(c). Has the benefit for the targeted group already been evaluated and reported in the research literature? If not, how will the assumed benefit be evaluated?

# P4: Scientific Integrity

P4(a). Is the impact of the nudge predictable, and based on evidence from the research literature?

P4(b). Does the chosen nudge (simple or hybrid) and the mode of delivery (once or repeatedly) match the decision (complex or simple) and behavior being targeted (short or long term)?

# P5: Social Responsibility

P5(a). Have the possible consequences of the nudge on the individual and society at large been considered? Have measures been undertaken beforehand to avoid or decrease possible negative side effects?

P5(b). Is there a reasonable plan for monitoring the e_ect of the nudge, i.e. taking snapshots at regular intervals?

P5(c). Is there a plan for discontinuing the nudge if unintended side effects are detected?

P5(d). Is there a proposal for monitoring long-term nudge impact if this is applicable?
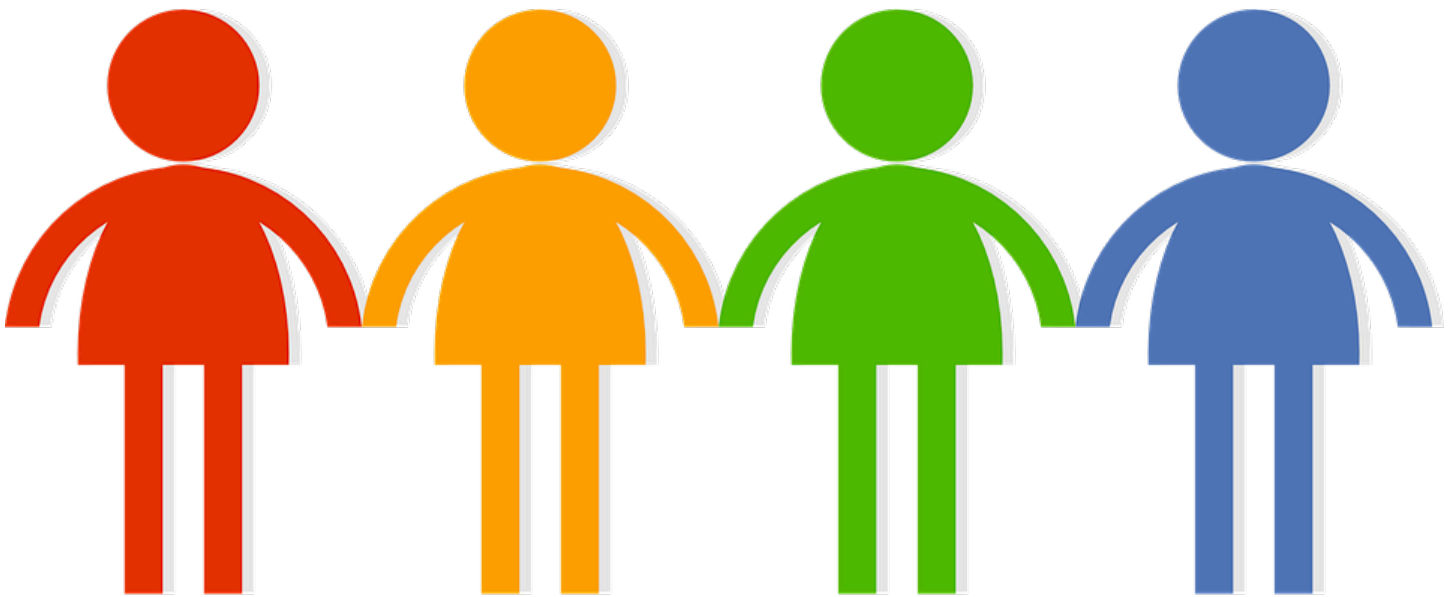
Renaud, K., & Zimmermann, V. (2018). Ethical guidelines for nudging in information security & privacy. *International Journal of Human-Computer Studies, 120,* 22-35.