

# Monitoring and Enforcement as a Second-Order Guidance Problem

Giovanni SILENO <sup>a,1</sup>, Alexander BOER <sup>b</sup> and Tom VAN ENGERS <sup>a,c</sup>

<sup>a</sup>*Informatics Institute, University of Amsterdam, the Netherlands*

<sup>b</sup>*KPMG, Amsterdam, the Netherlands*

<sup>c</sup>*Leibniz Institute, TNO/University of Amsterdam, the Netherlands*

**Abstract.** This paper aims to set up a conceptual framework for studying the *second-order guidance problem*—that is, designing coordination mechanisms for autonomous actors by means of adequate monitoring and enforcement measures—in a way which is sensible for designers and users of data-sharing infrastructures such as digital market-places. The paper outlines a minimal, but reusable and extensible computational model to test the sustainability of diverse norm implementations, evaluating it against relevant higher-level models presented in the literature.

**Keywords.** Monitoring, Enforcement, Reward, Punishment, Non-compliance, Policy design, Policy making.

## 1. Introduction

Data-sharing infrastructures as *digital market-places* (DMPs) manifestly exhibit the double status of computational and socio-economic systems.<sup>2</sup> On the one hand there are *physical constraints* on the operations that actors can execute; on the other, actors might entertain specific *contractual agreements*, there might be *market rules* and *societal norms* (e.g. GDPR and NIS directive) in place. Research and practice in DMPs are in general dominated by *control-oriented* views (on access and usage control, containment, security, ...). A gap exists in the literature in bridging between the control-oriented and guidance-oriented perspectives in a way which is sensible for designers and users of data-sharing infrastructures. The following example illustrates a possible application:

**Example 1 (Coordinating response to cyber-attacks).** *Consider a consortium of internet service providers (ISPs). One of the members is under cyber-attack. Information about the attack can be used to coordinate a collective defensive response, of which everybody will be eventually beneficiary. However, releasing information about the attack could provide access to competitive information. Certain parties might decide not to participate for reasons of economic opportunity. Which infrastructural policies should be implemented to promote the correct social functioning?*

Our goal here is to introduce a minimal, but reusable and extensible computational model to test the *sustainability* of certain monitoring and enforcement regimes, and their *effectiveness* with respect to given directives in a certain context.

<sup>1</sup>Corresponding Author: g.sileno@uva.nl.

<sup>2</sup>This paper results from work partly conducted for the NWO-funded project DL4LD (*Data Logistics for Logistics Data*, no. 628.001.001), and partly for the NWO-funded program VWDATA.

## 2. Modelling framework

*Norms expressions* One of the function of norms is to express *relative preferences* that should guide the behaviour of members of a society, typically of an action over its omission (or vice versa), or of the presence of a certain situation over its absence (or vice versa). This relative preference motivates the introduction of a norm, which can in turn be expressed in different ways. Two prototypical forms can be identified (cf. Hohfeld's framework of normative relationships): as a *deontic directive* (attributing a duty): "In context  $C$ ,  $X$  has the duty to  $A$ , otherwise she will obtain  $P$ ", or as a *potestative directive* (assigning a power): "In context  $C$ ,  $X$  has the power to obtain  $R$  by performing  $A$ ."  $P$  and  $R$  corresponds to two distinct enforcement regimes, based respectively on *punishments* (penalties, negative incentives, or the anecdotal "sticks"), and on *rewards* (positive incentives or "carrots"). Note that providing  $P$  and  $R$  to  $X$  requires the existence of some entity in the social system (typically some authority  $Y$ ) in the role of enforcer. Both directives can be rephrased without modality, from which we observe that  $P$  and  $R$  have formally the same role. So it seems that choosing between a carrot and stick regime is an arbitrary choice. But is it?

*Theoretical dimensions of norm application* Traditional approaches to enforcement take an internal view over the agent, typically based upon utility theory or other decision-making models. The introduction of a reward  $R$  and/or a punishment  $P$  typically modifies the expected value for the agent  $X$  associated to action  $A$ . Without enforcement, a rational constraint for deciding towards performance would have been:  $\mathbb{E}_X[A] > \mathbb{E}_X[\text{not } A]$ ; taking into account the enforcement we should consider:  $\mathbb{E}_X^R[A] > \mathbb{E}_X^P[\text{not } A]$ . (Our reference to utility theory here is just as an illustrative example of internal model). More importantly, in the following we will need to capture only the relative frequency of occurrence of conditions in which the agent  $X$ 's interests supersede the normative provisions. This measure, denoted with  $\text{PNC}_X$ , captures the **potential of non-compliance** of  $X$  for that norm.  $\text{PNC}_X$  is computed at individual level, but usually it is presented in aggregated forms, e.g. at population level, here denoted as  $\text{PNC}$ . It is a crucial *policy-field* value, required, even on a simplistic heuristic basis, to start discussions on any policy design.

Rather than looking at internal models, De Geest and Dari-Mattiacci [1] focus on the external dimensions of norm application, in particular compliance monitoring. Monitoring activity requires resources, and, depending on the context, monitoring for violation or for satisfaction might have different costs and probability of success. This makes the two regimes non-equivalent. As general considerations, the authors observe that sticks usually function better, but there are two cases in which carrots have to be preferred: in presence of a *specification problem* (difficulty of identifying the specific behaviour expected from the addressees); and of a *singling out problem* (non-uniform distribution of the burden over the addressees).

Rules about punishment and reward are conditionals. Someone needs to produce *evidence* of these conditions, even before (non-)compliance can be addressed. Boer [2] suggests that following the flow of evidence provides an alternative, even more essential way to look at the problem. In case of rewards, the agent claiming the reward is the one that has to provide the evidence; in case of punishment, it is the authority. This can be read as directly connected to a *default* rule (in reward-based enforcement, actors are deemed to be generally non-compliant, the opposite with punishment-based enforcement).

Authority	Agent X (addressee)	Collectivity
Monitoring cost: $m_p \cdot P(M) \cdot N$	Certification cost: $c_r$	Aggregated effects
Punishment benefit: $-p \cdot N_P$	Punishment cost: $p$	of performance:
Reward cost: $r \cdot N_R$	Reward benefit: $-r$	$(1 - \text{PNC}^e) \cdot P(C) \cdot N \cdot e_*$
	Non-normative effects	Aggregated effects
	of performance: $e_X$	of non-performance:
	Non-normative effects	$\text{PNC}^e \cdot P(C) \cdot N \cdot f_*$
	of non-performance: $f_X$	

**Table 1.** Economic voices distributed across different parties.

*Phases of normative interaction* We make here explicit the general phases associated to the operationalization of a norm: *performance*, *monitoring*, *enforcement*, and *certification*. Each phase can be played in principle by a different social actor, with different interests and view on the social system. The **applicability context**  $C$  in which the norm becomes relevant might have components which are static (e.g. spatially or temporally defined) and dynamic (agent behaviour, or environmental events); a distribution aspect can be observed at population level. Performance (or non-performance) by  $X$  can be motivated by reasons other than the norm, here captured by the condition  $C_X$ . This deliberation concerns only the **decision** to initiate performance ( $D$ ), whereas the outcome of the **action** ( $A$ ) might in general still be unsuccessful. We can define an *external non-compliance factor*  $\text{PNC}_X^e$ , including unsuccessful performances.

$$\text{PNC}_X^e = P_X(\text{not } A|C) = 1 - [1 - \text{PNC}_X] \cdot P_X(A|D)$$

Monitoring is the starting point for any enforcement. Typically it targets the presence of some **outcome** ( $O$ ) which is discriminant for the occurrence of the targeted action ( $A$ ), i.e. for which  $P(O|A) - P(O|\text{not } A) > 0$ . The quality of this discrimination can be captured by the posterior probabilities, namely  $P(\text{not } A|\text{not } O)$  (for punishment, in our running case) or  $P(A|O)$  (for reward). Informally, these probabilities measure a relative control of monitoring on the observation points relevant to action outcomes. We can distinguish two steps in **monitoring**: a *selection* mechanism, here captured with  $P(M)$ , and a *classification* step, whose predictive power for **violation** ( $V$ ) or **fulfillment** ( $F$ ) can be measured by  $P(\text{not } A|\text{not } O, M)$  or  $P(A|O, M)$ . In cases requiring stricter control, instead of advocating full surveillance ( $P(M) \sim 1$ ), a trusted third party (a *certifier*) could certify the action, improving the probability that the action is a proper one. For several reasons, there might be impediments to administering **reward** ( $R$ ) or **punishment** ( $P$ ); in the general case we should consider a probability  $P(R|F)$  or  $P(O|V)$ .

### 3. Sketch of economic flow

Monitoring and enforcement have a certain cost for the authority. All these costs will be eventually sustained by the social participants, according to some distribution (e.g. violators might contribute more through penalties). Let us denote with  $N_P$  and  $N_R$  number of punishments and rewards provided at runtime, with e.g.  $N_P = P(P|\text{not } A) \cdot \text{PNC}^e \cdot P(C) \cdot N$ ,  $N$  being the number of social participants; and with  $N_{VV}$  and  $N_{FF}$  respectively the actual (not the observed) numbers of violations and of fulfillments of the norm, which are proportional to the number of applicable cases  $N_C$ , following e.g.  $N_{VV} = \text{PNC}^e \cdot N_C$ ; in turn, the number of applicable and monitored cases is proportional to the population

$N$ , with  $N_C = P(C) \cdot N$ ,  $N_M = P(M) \cdot N$ . Let us assume that the corresponding collective effects are additive and grow linearly; we denote with the factors of growth  $e_*$  and  $f_*$  the *per-capita* distribution of the aggregated effects of all actions of performance (fulfillment) and non-performance (violation). For simplicity, we will neglect the monitoring cost voice for the authority in reward-based enforcement, but in this case we assume there might be certification costs on the addressee. Table 1 summarizes the distinct economic parameters for the different parties under these assumptions.

The *sustainability* of the system can then be captured by the following constraint:

$$(1 - \text{PNC}^e) \cdot e_* - \text{PNC}^e \cdot f_* \geq m_p \cdot \frac{P(M)}{P(C)} - p \cdot P(P|\text{not } A) \cdot \text{PNC}^e + r \cdot P(R|A) \cdot (1 - \text{PNC}^e)$$

This formula shows that attempting to bring PNC to 0 is in general not ideal, except perhaps for extremely critical contexts: besides reducing the space of autonomy for the social participants, the higher costs payed collectively might defeat the purpose.

From a theoretical point of view, it can be proven that the overall model (sustainability formula and internal model) confirms both De Geest and Dari-Mattiacci' [1] and Boer's [2] frameworks: (a) if people are generally compliant, too many "carrots" might easily make the system not sustainable; (b) reward-based enforcements become more effective if singling-out or specification problems are present; but also (c) when people are generally non-compliant. If performance is too expensive, avoidance becomes a rational choice, including contesting the authority. If *consensus* is part of the collective value structure, these effects, if quantified, would enter in the formula via  $f_*$ , eroding the surplus that was sustaining the punishment-based regime.

#### 4. Conclusion

The example presented in the introduction reflects two norms of an ISP consortium: (i) *If you suffer of a cyber-attack, share information with the consortium*; (ii) *If you are notified of a cyber-attack, start defensive maneuvers*. With adequate values for the environmental parameters, the proposed approach can be used to compute policy parameters for monitoring and enforcement. Clearly, the model presented here is simplistic and several assumptions are unrealistic. However, its extension is easy and straightforward, particularly in integrating e.g. capacities or other non-linear and circular<sup>3</sup> phenomena, non-additive relationships, sounder internal models, and various dynamical aspects, as for instance agents adapting to policies. We plan to extend the model in future study and investigate its application by means of (*optimization by*) *simulation* techniques, as our research targets on aspects of social-technical systems that cannot be treated by game-theoretical approaches based on static pay-off tables.

#### References

- [1] Gerrit De Geest and Giuseppe Dari-Mattiacci. The rise of carrots and the decline of sticks. *University of Chicago Law Review*, 80(1):341–392, 2013.
- [2] Alexander Boer. Punishments, rewards, and the production of evidence. In *Legal Knowledge and Information Systems Conference: JURIX 2014*, volume 271, pages 97–102. IOS Press, 2014.

<sup>3</sup>For instance, the probability to get a penalty in case of non-compliance is affected by the resources put in monitoring and enforcement. This probability in turn intervenes in the enforcement costs for the number of punishments and also influences the expectations of the agents, thus modifying the potential of non-compliance.