Recent years have seen the advent of the *microeconomic* paradigm in contemporary ICT systems design: hierarchical management architectures are giving way to federated system components which perceive mutual cooperation as beneficial. Numerous harbingers include distributed applications supporting collaborative work, multimedia content distribution systems, licensed spectrum sharing in LTE systems, social networks, online transaction portals, self-organizing systems and various novel architectures under the umbrella of the Future Internet. An abstraction of such systems is a community of intelligent agents interacting to exchange some quantifiable commodity called *service*. Examples of agents are resource operators, online entities, switching nodes, user terminals, smart devices, data centers, software robots, etc.; services can entail access to resources, information retrieval, participation in task execution, proxy assistance etc. A fundamental challenge arises from the contradiction between agents' qualities and expectations. On the one hand, agents are increasingly *autonomous* (i.e., act without external policing or supervision), *strategic* (i.e., weigh the benefits and costs of service exchange also anticipating other agents' behavior), and *privacy-concerned* (e.g., insist on some form of anonymity). In addition, because of their large number and/or privacy concerns, they may have little knowledge about agents being interacted with. Such features would ordinarily discourage any serious service exchange. On the other hand, agents increasingly depend upon one another as they collectively execute complex tasks and exchange sensitive data. A catalyst that keeps human agent communities going in the presence of this challenge is *trust*, defined as the extent to which an agent is willing to engage in a possibly risky interaction with another one. The idea is that agents who trust others are more likely to cooperate (provide more service) and so more likely to invoke trust in others – this is how trust precipitates *honest* behavior. A building block for trust is *reputation*, a perception of an agent's ability and willingness to cooperate created in others through past behavior.

Multiagent ICT systems interconnecting intelligent devices or online users are also interested in promoting honest behavior, hence digital equivalents of trust and reputation have to be developed. This mandates a dedicated functionality called a *reputation and trust building scheme* (RTBS). Application fields of such schemes are enormously vast: from collaborative interactive environments such as e-commerce to computer communication systems, such as cognitive radio and software defined networks, to e-governance. Notable examples include the CONFIDANT system for mobile ad hoc networks, opinion exchange fora in the Amazon, eBay and other portals; Google's PageRank webpage positioning can too pass as one. For all the multitude and diversity of existing or envisaged RTBS solutions, their systematic design is still considered in its infancy. Why should serious research be done into such schemes? The problem is that, if they are to serve truly commercial applications or even ones critical to systems security, one has to be clear how vulnerable they are to manipulation by *dishonest* agents and how to systematically achieve their main goal: promotion of *honest* behavior.

In a generic RTBS model, agents periodically select service providers to interact with, and next report the amount of received service (*reputation data*) to a *Reputation Aggregation Engine* (RAE). RAE computes agents' trust values and disseminates them among agents to support their future decisions regarding providing and reporting service. A dishonest agent wishing to manipulate RTBS in pursuit of a selfish or malicious agenda can launch a variety of attacks, e.g., to boost its own trust value while providing little service, or to boost the credibility of own reputation data or diminish that of other agents. Some attacks are more damaging and harder to defend against in the case of attackers' *collusion*. At present, methods of systematic RTBS design are sought, in particular ways to make an RTBS resistant to strategic dishonest agents. This is considered a soft security matter, where the traditional hard security paradigm of preventing access by dishonest agents is supplanted by one of merely discouraging their misbehavior and reducing the damage it causes.

RTBS design becomes more difficult when selection of service providers is "blind" – forced by current availability of particular services or (especially in mobile environments) geographical proximity, and not dictated by trust values. This frequent situation is rarely accounted for and calls for novel design methods. The only defense of honest agents then lies in the differentiation of their behavior depending on the interacting agent's current trust value, that is, rewarding its honesty toward third-party agents according to the principle known in biology and sociology as *indirect reciprocity*. Within the project, analytical and simulation studies of ICT systems under agents' anonymity and "blind" selection of service providers will be conducted using interdisciplinary methods to answer a number of fundamental questions, such as: Can dishonest agents acquire higher trust values than honest agents (if so, RTBS loses its point since high trust values are no longer in demand)? What should be done to avoid such as scenario? What is the impact of the proportion of dishonest agents and the presence or absence of their collusion? What amount of received service can honest agents count upon in the presence of selfish or malicious dishonest agents? Is indirect reciprocity beneficial? Would it be beneficial if the algorithm whereby RAE arrives at trust values were concealed from agents?