



## Concordance in the Crowdsourcing Activity

Viacheslav E. Wolfengagen<sup>1</sup>, Larisa Yu. Ismailova<sup>1</sup>, and Sergey Kosikov<sup>2</sup>

<sup>1</sup> National Research Nuclear University “MEPhI”  
(Moscow Engineering Physics Institute),  
Moscow, 115409 RF,

{jir.vew@gmail.com; lyu.ismailova@gmail.com}

<sup>2</sup> Institute for Contemporary Education “JurInfoR-MGU”,  
Moscow, 119435 RF,  
kosikov.s.v@gmail.com

### Abstract

A concordance in cognition activity of possibly interrelated crowdsourcers aimed to property recognition in the voluminous data sources is considered. Data sources are of either usual nature or manually generated with the crowdsourcing. The proposed model is based on the variable domains assumption. A general layout is able to take into account an interaction of crowdsourcers and properties when they are varying with the evolving the events. The cognition model is of stage-by-stage type and has the representable functor. This model as may be shown is faithfully embedded into a category of indexed sets. Using the proposed neighborhood for cognition activity leads to a flexible computing model.

*Keywords:* crowdsourcing, Big Data, Thick Data, variable domains, cognition model, concordance, computational model

## 1 Introduction

There are no commonly adopted in use cognitive architecture and model of cognition. In this paper we build the model of cognition based on assumption of evolving the knowledge “stages”. This paper is inspired by a some kind of similarity between the “natural” [5] and “ordinary” computing which we mostly are familiar with. A generic notion for both kinds of computing is the *information process*. Here is an attempt to construct its model to cover the type-generating procedure which is fundamental for information processes. To achieve the aim a kind of homotopy based model is involved and applied to an area of crowdsourcing and the practical activity of crowdsourcers. In their manual processing and generating the diverse data sources they often work as a kind of “knowledge engineers”. And this activity needs a faithful computation model based on rigorous “knowledge mechanics”. This is done using the models relating to applicative computing systems [1] where the main point is the *interplay*, or interaction of the objects of a rather abstract nature.

A task of property recognition has a vital importance in analysis of Big Data to extract the needed in practice Thick Data [4]. The modern trend in solving this task is to attract a group of crowdsourcers. The technology of crowdsourcing generates, in addition to usual sources of data, manually produced data sources which have a trend to tremendous growing.

Such a “thickening” of data is the main point of a *type theory*, but to bring in this process more flexibility some key ideas of a homotopy type theory [9] will be used. As a consequence, this will allow to *explicitly* include some extra-parameters to characterize the action of a group of, possibly, interrelated crowdsourcers.

The interconnections of social media, data integration, and human computation can be found out in [2]. A confidence of the crowd was analyzed in [8]. Variable domains in a category theory were refined in [9].

A preliminary study of possible media of interaction was in [11] for semantic nets and in [10], [12] for applicative models in case of interaction of objects with environment. The vision of information processes from a point of view of computational thinking was studied in [6], [7].

In Section 2 there is a characterization of a property recognition with a group of experts/crowdsourcers. They are equipped with a stage-by-stage knowledge acquisition scheme (cognition model). Section 3 contains a construction of representative category which gives a faithful basis to proposed  $(f, g)$ -neighborhood model and corresponding knowledge acquisition media. This model gives an embedding for an interaction of crowdsourcers and properties which they are recognizing.

## 2 Crowdsourcers and the property recognition

In short, crowdsourcing can be perceived as the control of assignments and properties/features which arise in analysis of Big Data. This needs a series of analytical actions including: selecting out the collections for extracting the information in need; extraction of entities from the non-structured or weak-structured sources; evaluation of neighborhood of individuals in different data collections; merging the instances of individuals including detection and erasing the duplicates; development of data schemes in collections and their mapping into the target scheme; generation of instances of the actual individuals and concepts for data corresponding the target scheme.

### 2.1 The technology of crowdsourcing.

The area of crowdsourcing is extremely booming, and so far the conventional systems and methods of working with them has not yet appeared. With crowdsourcing [3] the handwork is used for data processing, obtaining or generating data on demand, as well as for their classification, ranking, marking or refining the existing data. These manually solvable problems are often difficult to automate, for example, when determining the ranking of something or someone, or in determining any signs of interest to any data source. The data, manually created, can also be viewed as a source of data on equal rights, so naturally, we would like to integrate this crowdsourced data with other traditional sources. This will allow the end user, instead of working with heterogeneous data sources, to communicate with a single unified database, which is an advantage.

### 2.2 A neighborhood of the property.

A search of neighborhood of the property is depicted in Figure 1.

Download English Version:

<https://daneshyari.com/en/article/4962287>

Download Persian Version:

<https://daneshyari.com/article/4962287>

[Daneshyari.com](https://daneshyari.com)