



Notice

Date

July 15th, 2016

Subject

Market survey automatic field sketches reading

From

Ad Spaan

Contact details

inkoopservices@kadaster.nl

Page

1 of 10

To the interested party,

The Dutch Cadastre (from here on called Kadaster) is hereby inviting you to take part in a market survey.

With this market survey we hope to collect information about the feasibility of automatic reading and interpreting of analogue (*.jpg) technical drawings. In this case these are the so-called field sketches in which the surveyor has written down dimensions that describe the exact location of parcel boundaries. Because of the enormous number of fields ketches (more than 10 million) interpreting these documents by hand is expected to be far too expensive.

You have the possibility to address additional questions until the 29th of August 2016. All correspondence related to this market survey should be addressed to the email-address listed below.

Sincerely,

Ad Spaan

Email: inkoopservices@kadaster.nl



Table of contents

1	Market survey: automatic field sketches reading	3
1.1	Introduction.....	3
1.2	Goal of the Market survey	3
2	Organisation & project	4
2.1	About Kadaster.....	4
2.2	Background	4
2.3	Project description	4
2.4	Process and goal.....	5
2.5	Goals in scope.....	5
2.6	Out of scope	6
3	Test case (example field sketch) and rules	6
3.1	Test case.....	6
3.2	Rules	6
3.3	Measurements.....	6
3.4	Possible process	6
3.5	Additional material	7
4	Procedure.....	7
4.1	Market survey schedule.....	7
5	Overview of documents in the package (supplements)	7
	Appendix A – Questions and answers	8
	Appendix B – Template for Questions	10
	Appendix C – References.....	10

1 Market survey: automatic field sketches reading

1.1 Introduction

More and more, house and landowners want to have access to legal information of their own property. Currently they are very dependent on Kadaster and have to pay to get the information. The cadastral map, which is used in this process, is designed as a geographical view on the registration with limited geometric quality and it fits that purpose very well. However, people will zoom in and rely on the details they find. Due to the nature of the data, a distance between two lines might have errors up to half a metre on the map. These differences may lead to misunderstandings and wrong assumptions. Kadaster wants to improve the communication about the cadastral map but also to enhance the quality of the map. For that purpose, we started a research project to find out the possibility to create this improved map using automated processes. We intend to do this by interpreting the information found on field sketches.

In this market survey we are asking companies to provide us with information that might contribute to partial or complete solutions to our very specific questions. The keyword here is feasibility, which is a key indicator in our decision making process. We are providing a case with a description of the challenge and a description of a possible outcome of the process. We also provide you with information about the rules that define the content of the field sketches. Interpreting the information found on the field sketches is our primary goal for the initial step in the process.

1.2 Goal of the Market survey

The goal of this market survey is to get more insight in:

- which solutions are available for automatic conversion of field sketches as a whole or as partial solutions?
- which companies offer these solutions?
- what are the costs involved with implementing the proposed solutions?

The results of the market survey will be used to get a better insight in technological developments of conversion of scanned documents to usable digital information. These results help Kadaster choosing a direction and determining the right selection criteria for a possible purchasing process.

All supplied information will be treated confidential, although a summary of the proposed solutions will be shared. This information and the other results of this market survey will not be used to select a supplier, nor will it ensure any future follow-up orders. There will be no payment for the supplied results.

2 Organisation & project

2.1 About Kadaster

The Netherlands' Cadastre, Land Registry and Mapping Agency – in short Kadaster – collects and registers administrative and spatial data on property and the rights involved for real estate, utility and telecom networks, ships and aircraft. In doing so, Kadaster protects legal certainty. Kadaster is also responsible for national mapping and maintenance of the national reference coordinate system. Furthermore, Kadaster is an advisory body for land-use issues and national spatial data infrastructures and has an international office that coordinates the international consulting activities of Kadaster.

2.2 Background

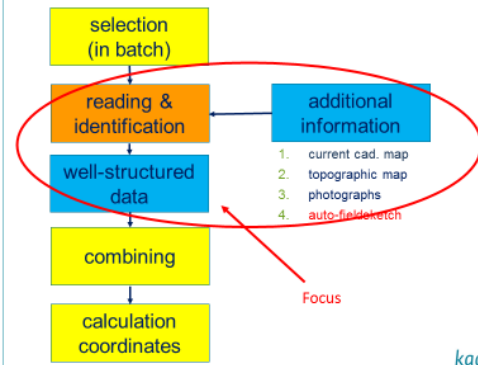
Kadaster observes an increasing demand for information of people in general and more specifically of citizens requiring information about their real estate. The description of parcel boundaries is important for landowners in case of disputes with neighbours. People also want to have a clear insight in the size (in square metres) of their parcel. The cadastral map was started by Kadaster as a geographical index for parcels and, even though it is not intended to be very accurate, also used for calculating parcel sizes. For more precise and accurate information regarding parcel boundaries people have to consult reconstruction information. This information is stored on field sketches which contain detailed data of small areas. Due to our legal foundations the location of each parcel boundary is defined by the original data recorded at the time of the creation of the boundary. However, since more and more people are using the more easily available cadastral map to derive boundary information, the need for a more accurate cadastral map is increasing.

2.3 Project description

Kadaster aims to improve communication about the quality of the cadastral map and to improve the quality of the map itself by bringing the reconstruction precision to the map. For this second goal a research project is started. One of the results of this project will be a map with more accurate coordinates for 90% of the boundaries. For 5 to 10% we lack the information to improve the map. For information which is already stored in digital coordinates this is nearly 100% possible. For about 60% of the parcels however, data is located on scans of field sketches (analogue drawings).

Using the dimensions on these sketches, it is possible to calculate coordinates of the boundaries. Manually reading this will be very time consuming and therefore very expensive. To cope with this challenge we are looking for possibilities to automate the reading process. This part of the challenge is the question we are asking the market. More specifically, we are interested in the feasibility of a solution to read and interpret our field sketches, and an estimation of the percentage that can be read.

Production of improved map Proces of calculation of new map coordinates



2.4 Process and goal

In the picture the data extracting and calculation process is shown. As marked the focus of this market survey is on the reading and identification process with the help of reference data (additional information). In the included data-package a much more comprehensive scheme can be found. The goal is to deliver well-structured data in a way that is as automatic as possible.

2.5 Goals in scope

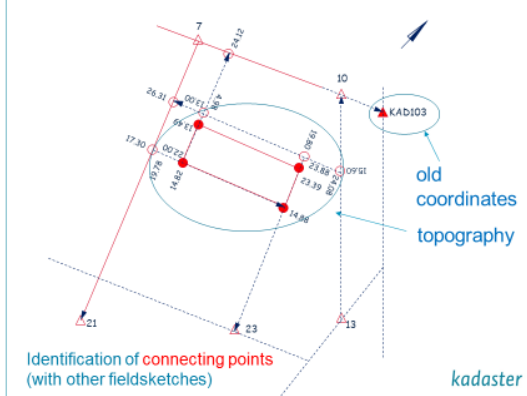
Our specific goals are to have a good estimation of the feasibility of:

1. the reading of analogue field sketches:
 - a. The handwritten text and numbers, together with the writing direction
 - b. The vectors, intersection points and symbols
2. The interpretation (recognition and classification) of the information that is read in accordance to the rules of the field sketches construction:
 - a. The determination of colour, texture and shape of elements
 - b. The combination of the information into points, measurements and additional information
3. Integrating reference data into the search process

We are also interested (but only secondary!) in suggestions to solve the problem of:

4. Determining connecting points that will help us to interconnect floating coordinate fields from different field sketches (see picture and Supplement 4)
5. Reading with OCR handwritten coordinate lists and extracting the relevant information out of it (see Supplement 8)

Extracted information of fieldsketch



2.6 Out of scope

There are also challenges that we willingly leave outside this market survey. The whole data handling process like selecting the field sketches, the combining process of “floating coordinate fields” and the calculation process itself are left out. We already have done some testing and have the capability of solving these processes ourselves (although this does contain a lot of work).

3 Test case (example field sketch) and rules

3.1 Test case

In our package of supplements you can find an example of a field sketch in the Dutch town of Assen, created in the 70's. We kindly ask you to take this example and try to reach the described goals. In the package you will also find an example of well-structured data from which we can take off to start the calculation process with the MOVE3-application. We used rather complex point numbers in the example with information of the location of the points (like parcel numbers and house address numbers) stored in it. If you like you can use simple point numbers as long as they are unique within this example. The complexity of this field sketch is average, in the package you will also find some other examples with different complexity.

3.2 Rules

We provide you with a very expansive set of rules. In general the used measurement method has been the same throughout the years, namely intersection of lines between known (in coordinates) points and lines in extension of building walls, in combination with measured distances. The measurements are always in meters and nearly always with two digits after the decimal point (which mean cm's). These numbers are written in a direction perpendicular to the measuring direction and readable from the zero-position. The zero position is marked with an arrow. The number is written as close to the intersection point as possible and on the most suitable side of the line (depending on the available space on the paper it can be right or left). More details and exceptions are described in the supplement, much more than needed to solve the test case.

3.3 Measurements

In the test case three types of measurements (=relations between two or three points) can be found: tape distance (TD), collinearity (CL) and perpendicular (PD). You can find the explanation in the included structured data example (Supplement 2). For calculation purposes approximate coordinates of all points are needed, they can be calculated by our calculation program but also extracted from a map in advance.

3.4 Possible process

We designed already some schemes in which we try to solve the problem theoretically. Due to a lack of experience in the field of deep learning systems we cannot say that this system is working, but one might use this as a way to visualise all the aspects involved in the challenge. We are open for suggestions to change the chosen solution direction.

3.5 Additional material

Within the package a lot of additional material can be found: examples of other field sketches with differences in complexity, age and region. Some of the rules relate to regional differences in the field sketches. You can use this material to get a better impression of our challenge.

4 Procedure

To facilitate the provision of answers to the questions raised by Kadaster for the supplier, please use the available format: Appendix A - questions and answers. If you have questions about (parts) of the market survey, we ask you to use Appendix B – Template for Questions.

We also ask you to supply two (2) references to which is a similar product or service has been delivered. Please use Appendix C – References for this.

Based on the information submitted Kadaster will determine the right selection criteria for a possible purchasing process.

4.1 Market survey schedule

This market survey will be placed on TenderNed on the 1st of August 2016. The closing date for the survey is the 9th of September 2016.

Questions can be asked until the 29th of August 2016 by sending an email to inkoopservices@kadaster.nl. Answers (together with the questions) will also be published on TenderNed.

5 Overview of documents in the package (supplements)

Together with this document a package is sent which contains the following documents.

Test case

1. Test case field sketch of average complexity
2. Example of structured data (excel files) extracted from the test case field sketch
3. Drawings that show the inputs and desired output of different steps

Documentation

4. Information regarding connected information (neighbour) field sketches
5. Schemes that show the possible process of reading, analysing and combining data

Examples showing the complexity of the challenge

6. Legend: rules glossary and exceptions related to different regions
7. Examples of other field sketches (4 regions)
8. Old-fashioned lists of ground points and its coordinates

Appendix A – Questions and answers

#	Question	Answer
Your company and product		
1.1	Clarify your company, vision and strategy. Supply key figures regarding your company.	
1.2	Is the proposed solution an existing product or service, or will it be developed? Are there any clients who already use your solution? How many clients are using it? For how long they have been using it? What is the feedback that you take from them?	<i>Add two references using Appendix C.</i>
Proposed solution		
2.1	What is the complete or partial solution that is proposed to read and interpret information on the field sketches and convert it to a usable digital format?	<i>Please take more space to demonstrate your solution.</i>
2.2	What steps are identified in the process? Also specify, who is responsible for these steps. If Kadaster has a role in specific steps, please indicate this.	<i>Please take more space to demonstrate your solution.</i>
2.3	What tooling / software are you using for the implementation of the solution? Is this tooling / software developed by your company or is it commercially available on the market?	
2.4	Information on the field sketches is written in the Dutch language. Does your solution / product / service have the capability to interpret the information correctly?	
2.5	What is the estimated time required to read and interpret per sketch?	
2.6	What is the amount of sketches that can be read and interpreted per day / week? How easy is it for you to upscale this process?	
2.7	How will you guarantee the quality of your output? Which method will you use for that?	
2.8	Do you have any additional requirements for Kadaster regarding the source data?	
2.9	Do you have any additional requirements for Kadaster regarding the IT infrastructure?	
Technical questions		
3.1	In some field sketches the measurements are written without a decimal separator. Sometimes the decimal	

	part is written as superscript and sometimes there is a small line below the decimal part. How will you ensure the measurements are interpreted including the decimal part?	
3.2	In some cases the location of the measurement text might be confusing (e.g. not close to the point/line that is being referred to). Describe how the measurements are connected with the correct points.	
3.3	Some points are both building outline points and parcel boundary points (when a building line is located on a boundary). We want to have both and defined as identical. How will you handle these cases?	
3.4	Some field sketches contain measurements from reference points which are not included in the observed field sketch but in a connected one. The information to determine which connected sketch can be found on the sketch. How will you handle these cases?	
3.5	How will you deal with situations where the information on the field sketch is incomplete?	
3.6	Indicate where you need further explanation, based on the included field sketch samples.	
Costs		
4.1	What can you tell us about the expected costs, either for proposed partial solutions, or the solution as a whole?	
4.2	What is your cost structure and where are the biggest costs expected? Make a distinction in costs related to: initial investment, start-up costs, usage and maintenance costs.	
4.3	Do you use a fixed price structure, time & material, or other?	

Appendix B – Template for Questions

#	Page / chapter / paragraph	Question
1		
2		
3		

Appendix C – References

Reference details	
Company name	
Address	
Contact person	
Job position of contact	
Telephone number	
Industry sector	
General details	
Project start date	
Project end date	
Project size in € (excl. VAT)	
Project details	
Project title	<i>Specify the name or title of the project.</i>
Project size	<i>Specify the number of documents read and interpreted.</i>
Approach	<i>Specify the project approach (use max. 1 A4). Indicate which part was executed by your company and which part was handled by subcontractors.</i>
Tooling / software used	<i>Specify what tooling and / or software was used in the project. Indicate what was your own tooling / software and what was not.</i>
Size of project team	<i>Specify the size of the project team for your own company, the reference organization and for subcontractors if any.</i>
Project outcome	<i>Describe the project outcome and results.</i>