



## Bachelor project Computing Science – Starting form

*To be filled in and signed by the student and the supervisor(s) at the beginning of the bachelor project.*

Course code: WBCS13000

Credits (EC): 15 points (420 hours)

*Please note that at the start of the Bachelor project the student should have completed the propaedeutic phase and earned 80 ECTS from year 2 and 3*

---

### 1. General information

Student name and signature: Matej Kucera

Student number: s4551192

Project title: Seal photo-identification database conversion

Date: 20/04/2023

Starting meeting: 17.4.2023

Final assessment: 7.7.2023

First supervisor (name + signature): Andrea Capiluppi

Second supervisor (name + signature): Beatriz Rapado-Tamarit | Zeehondencentrum

Daily supervisor (if applicable):

External supervisor (if applicable):

---

### 2. Project description:

The stated problem is that the researchers at Zeehondencentrum Pieterburen have been using the ExtractCompare software to photo-identify seals. They have two large databases. However, they now want to merge the two databases together, which is impossible to do in ExtractCompare. Hence, this project will look at the options of converting the database to different software with the same functionality which will allow the databases to be merged.

The best candidate seems to be IBEIS, a software which satisfies all of the criteria. This was found during the project proposal stage. It was also found that there are many other applications which have similar functionality in case IBEIS does not work out.

The plan is to take the MS Access based databases which currently store the data and write a python program which can read them and automatically insert them into IBEIS using the API IBEIS provides. There is a possibility that the data may have to be



altered or otherwise modified, which will be done as necessary. Finally, importance will be placed upon testing. This is to ensure quality of resulting code and correctness. Documentation will be provided with the resulting application as well to ensure that researchers with little to no technical skills are able to use the application with ease.

**RQs:**

Which software is the best at the moment to photo-identify seals?

What is the best way to convert an ExtractCompare based seal image database to another database type?

What does the conversion need to do in order to enable the merging of two databases?

**3. Methodology and timeline:**

IBEIS is written in Python and exposes an API which can be used to access its data management and algorithms. This can be used to automatically import all the seal images. A library for accessing MS Access databases exists for Python called pyodbc. This library can be used to read the ExtractCompare database and get all the information from it. Then, the information can be fed to the IBEIS API to create a new database. If the API requires additional information about the picture such as the position of the animal, this will be implemented using machine vision in python using machine learning libraries like pytorch, or a ready-made solution if one can be found.

Testing will be set up in a way that will verify the resulting converted database in several ways, including integrity of data, accuracy of matching and ease of use of both IBEIS with the new database as well as the conversion tool itself. Heavy focus will be placed on testing to prove that the conversion works smoothly and without errors.

The application will be written in Python due to ease of use and compatibility with IBEIS. The databases to be converted for Zeehondencentrum Pieterburen are around 22 GB total. These are stored both in the seal centre as well as on the student's computer. To transfer them, an encrypted email based service was used. The database will not be made public and will not be stored under version control. Backups of the database have been made to ensure no information is lost in case of an error, since it is not under version control. For version control of the software itself, a repository on GitHub will be created where the code will be stored. This repository will be private during development and possibly published when it is finished.

The milestones of the project, planning and deadlines can be seen in the table below.



Dates	Plan	Deliverable
17.4	Project start	
18.4	Planning	Detailed plan
19.4 - 21.4	Read IBEIS API	notes on API usage
24.4 - 26.4	Explore seal database	Database schema diagram
27.4 - 28.4	Read DB into Python	Code
1.5 - 2.5	Test database	Tests, correct database
3.5 - 17.5	Connect app to API	Code
18.5 - 26.5	Test DB conversion & IBEIS	Tests
29.5 - 2.6	Convert and merge DBs	IBEIS database, code
5.6 - 9.6	Test IBEIS database	Testing
12.6 - 13.6	Documentation	Documentation
14.6 - 23.6	Write thesis	Thesis
28.6 - 30.6	Presentation	Slides + Live presentation
3.7 - 7.7	Finish up all deliverables	Submission

#### 4. Division of tasks

Not applicable.

#### 5. Deliverables:

- Software source code with tests
- Documentation for the software
- Database (converted and merged from the two databases)
- Thesis
- Presentation

#### 6. Grading

Scientific quality of research and technical contributions: 40%

Project management and interpersonal skills: 20%

Final Presentation: 20%

Report/Thesis: 20%