

**Y1A 2021**

Project Brief

Academic year: 2021-2022

Educational year: Year 1

Block: Block A

Credits: 15 ECTS

Osiris Code:

Last update:

Contact:

Contents

[1 Project Brief 3](#_Toc72233233)

[1.1 Brief Outline 3](#_Toc72233234)

[1.2 Project Kick-Off 3](#_Toc72233235)

[1.3 Competency Profile and Dublin descriptor 3](#_Toc72233236)

[1.3.1 Competency 4: Research & Analysis Replace 3](#_Toc72233237)

[1.3.2 Competency 5: Conceptualizing 3](#_Toc72233238)

[1.4 Learning Objectives 3](#_Toc72233239)

[1.4.1 Personal Development & Academic Practice 3](#_Toc72233240)

[1.4.2 Professional Practice 3](#_Toc72233241)

[1.4.3 Artificial Intelligence and Digital Transformation 4](#_Toc72233242)

[1.4.4 Data Science 4](#_Toc72233243)

[1.4.5 Programming 4](#_Toc72233244)

[2 Creative Brief I - Report on AI in Science Fiction 5](#_Toc72233245)

[2.1 Timeline 5](#_Toc72233246)

[2.1.1 Calendar week 1 5](#_Toc72233247)

[2.1.2 Calendar week 2 5](#_Toc72233248)

[2.2 Assignment 5](#_Toc72233249)

[2.3 Requirements 6](#_Toc72233250)

[2.4 Deliverable(s) 6](#_Toc72233251)

[2.5 References 6](#_Toc72233252)

[3 Creative Brief II – Poster: Our world in data 7](#_Toc72233253)

[3.1 Timeline 7](#_Toc72233254)

[3.1.1 Week 3 7](#_Toc72233255)

[3.1.2 Week 4 7](#_Toc72233256)

[3.1.3 Week 5 7](#_Toc72233257)

[3.2 Assignment: Poster-Our world in data 8](#_Toc72233258)

[3.3 Challenges 9](#_Toc72233259)

[3.4 Requirements 10](#_Toc72233260)

[3.5 Deliverable 10](#_Toc72233261)

[4 Creative Brief III – Image Data processing using python 12](#_Toc72233262)

[4.1 Timeline 12](#_Toc72233263)

[4.1.1 Week 6 12](#_Toc72233264)

[4.1.2 Week 7 12](#_Toc72233265)

[4.1.3 Week 8 12](#_Toc72233266)

[4.2 Assignment: Image Steganography using Python 13](#_Toc72233267)

[4.3 Challenges 13](#_Toc72233268)

[4.4 Requirements 14](#_Toc72233269)

[4.5 Deliverable 14](#_Toc72233270)

[4.6 References 14](#_Toc72233271)

# Project Brief

Soft skills / personal development / learning log / work log / Assessment sheet and it’s purpose

Where do we put info on our teaching philosophy/relation GameLab-Github etc.

A creative brief is like an open invitation. It invites an individual to propose their own creative vision while still giving a clear direction on the desired outcomes. It is quite customary in the data science industry for projects to have abstract requirements. A typical example being “Can we improve our existing business processes using data?”

Similarly, our creative briefs set out to align with specific learning goals. The requirements set in the brief below will lead junior data scientists to learn about different topics, while still allowing for creative freedom in proposing and implementing a solution. Select an approach that best fits your personal learning goals, investigate the requirements, and amaze us with your creative problem solving!

Challenges bronze-silver-gold

## Brief Outline

In this project, you will be exploring the *lifecycle of a data science project* by implementing in practice, the concepts you have learned in the workshops. You will be provided with a Github repository where you are expected to document the process of forming your idea, as well as the evolution of the solution.  The project consists of three sub-projects, the first one related to Digital Transformation and Artificial Intelligence, the second one to Data Science and the final one to Programming.

## Project Kick-Off

In the first week of the project, you will have time to meet with each of the lecturers, who will each take time to discuss and direct one aspect of the project. The aim in this time is to provide a strong, solid foundation for the rest of the project – so pay attention carefully to the feedback and direction given and execute it as efficiently and accurately as you can. This period of work is not covered by the assessment but will heavily affect how easy it is for you (and the rest of the team) to work towards the assessment later.

Procedures

## Competency Profile and Dublin descriptor

### Competency **4:** Research & Analysis Replace

The starter professional practitioner is capable of substantiating a design assignment by means of research and analysis. The starter professional practitioner demonstrates in his/her research activities that he/she has a repertoire of relevant research skills and is able to select the correct method from this repertoire, given the research circumstances. Is capable of developing prototypes as a means of communication within the context of the application.

Dublin Descriptors: Knowledge & Insight, application of knowledge & Insight, Making Judgement

### 1.3.2 Competency **5:** Conceptualizing

The starter professional practitioner demonstrates he/she is capable of achieving a realistic cross-sector demand articulation and project definition. He/she is capable of developing, on the basis of his/her own idea or demand articulation, an innovative concept which creates value.

Dublin Descriptors: Knowledge & Insight, application of knowledge & Insight, Making Judgement

## Learning Objectives

### Personal Development & Academic Practice

Demonstrate self-exploration and personal development, good academic practices in learning how to learn and the acquisition of professional knowledge through research, study, analysis, applied practice, discussion and reporting.

### Professional Practice

Demonstrate professional behavior as well as accountability and ethics in the application of industry best practices for planning, communication, collaboration, and responsible execution of work assignments.

### Artificial Intelligence and Digital Transformation

The student demonstrates the basic knowledge about the field of AI and how AI is applied in different sectors of the economy by analysing its foundational concepts and reporting current cases in which AI is applied to transform companies by means of an effective (informative, readable, clear) report.

### Data Science

The student demonstrates a full understanding of different data types, descriptive data, graphical representations of data, statistical inference, basic probability theory, correlation and simple linear regression by solving appropriate use-cases. The student is able to utilize different data visualization methods available and choose the appropriate graphical representation to gain insight from data and interpret the visualisations appropriately. Report results in an insightful manner.

### Programming

Students develop a basic understanding of python programming concepts, data types and data structures and can apply learned concepts and use standard python libraries to solve use-cases.

# Creative Brief I - Report on AI in Science Fiction

## Timeline

### Calendar week 1

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Day/Time | 9-10 | 10-11 | 11-12 | 13-14 | 14-15 | 15-16 | 16-17 |
| Monday (self-study 1) | Foundations of AI (1): Philosophy, and history of AI  Turing Test & Chatbots (1): Article and chatbots | | | | | | Q&A and day recap |
| Tuesday (DataLab 1) | AI in Science Fiction (1): Minority Report screening  Turing Test & Chatbots (2): Discussion and experiment | | | | | | Day reflection |
| Wednesday (Self-study 2) | Reporting: The structure of a paragraph and rules of writing Reviewing and re-writing a paragraph  Citations and bias-free language  Mechanics of style of effective writing | | | | | | Q&A and day recap |
| Thursday (Self-study 3) | Foundations of AI (2): Symbolic/Non-symbolic AI  Intelligent agents | | | | | | Q&A and day recap |
| Friday (DataLab 2) | AI in Science Fiction (2): Formative feedback session on first draft | | | | | | Day reflection |

### Calendar week 2

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Day/Time | 9-10 | 10-11 | 11-12 | 13-14 | 14-15 | 15-16 | 16-17 |
| Monday (self-study 1) | AI for Business (1): Introduction into application of AI in businesses  Risks and benefits of AI | | | | | | Q&A and day recap |
| Tuesday (DataLab 1) | AI in Science Fiction (3): Q&A session on the final report | | | | | | Day reflection |
| Wednesday (Self-study 2) | AI for Business (2): Application of AI in businesses | | | | | | Q&A and day recap |
| Thursday (Self-study 3) | AI for Business (3): Application of AI in businesses | | | | | | Q&A and day recap |
| Friday (DataLab 2) | AI in Science Fiction (4): Q&A session on the final report | | | | | | Day reflection |

## Assignment

Students are required to create a professional report of maximum 2000 words (excl. cover/table of content/list of references and appendices) on AI and Science Fiction using academic style of writing. For the report, students are expected to analyse Minority Report, which is a movie that uses AI as an integral part of the plot.

In the report, the student must describe the plot and the role played by AI; analyse the author(s)’ concept of AI and conclude by discussing the feasibility of this concept. Additionally, the report needs to contain a novel case of how this, or a similar AI concept as in the movie can applied in a company, business, or organization and solve its problem. Lastly, the risks and benefits associated with this AI application need to be examined and described.



Figure 1. Minority Report (Spielberg, 2002)

## Requirements

* Maximum 2000 words (excl. cover/table of content/list of references and appendices)
* Identify, and describe an AI concept within the movie Minority Report
* Explain how this AI concept or a similar AI concept as in the movie is being implemented in 3 business of your choice
* Find a business that has a challenge that can be solved by applying this AI concept
* Explain how this AI concept can be used to solve the business challenge of this business
* List at least one benefit, and one risk associated with this specific AI business application
* Effectively read and summarize (academic) literature
* Support an argument by providing evidence found in relevant (academic) literature and theory
* Use proper English language, structured and functional layout and properly use references according to APA style

## Deliverable(s)

* Formative: Presentation: cite its dominant AI theme(s), and outline the review
* Summative: Report on AI and Science Fiction

## References

# Creative Brief II – Poster: Our world in data

## Timeline

### Week 3

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Day/Time | 9-10 | 10-11 | 11-12 | 13-14 | 14-15 | 15-16 | 16-17 |
| Monday | 1. Introduction to DS (seeing the world in data, attributes) 2. Interactive Workshop and introduction to R using swirl. 3. Interactive Mock Assessment | | | | | | Q&A and day recap |
| Tuesday | DataLab 00: Introduction.  Choose SDG indicator (<https://sdg-tracker.org/> ), create problem statement and setup introduction | | | | | | Day reflection |
| Wednesday | 1. Variables (data frames, continuous, nominal, ordinal etc.) 2. Interactive Workshop 3. Interactive Mock Assessment | | | | | | Q&A and day recap |
| Thursday | 1. Descriptive analyses (mean, sd, range, IQR) & visualisation (boxplots) 2. Interactive Workshop 3. Interactive Mock Assessment | | | | | | Q&A and day recap |
| Friday | DataLab 01: Exploratory Data Analyses (EDA) Methodology  Identify and download the relevant SDG variables. Import the dataset and packages. Plan and perform an exploratory data analysis and report your methodology. | | | | | | Day reflection |

### Week 4

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Day/Time | 9-10 | 10-11 | 11-12 | 13-14 | 14-15 | 15-16 | 16-17 |
| Monday | 1. Introduction to Probability (random variable, distributions) 2. Interactive Workshop 3. Interactive Mock Assessment | | | | | | Q&A and day recap |
| Tuesday | DataLab 02: Findings  Create informative descriptive data visualisations from your EDA. | | | | | | Day reflection |
| Wednesday | 1. Introduction to stat. inference (sample, pop, hypothesis testing) 2. Interactive Workshop 3. Interactive Mock Assessment | | | | | | Q&A and day recap |
| Thursday | 1. Analysing Relationships between variables (e.g., compute correlation by hand) 2. Interactive Workshop 3. Interactive Mock Assessment | | | | | | Q&A and day recap |
| Friday | DataLab 03: Discussion  Identify and write up the interpretation considerations regarding the data and make recommendations for future analyses and data driven decisions. | | | | | | Day reflection |

### Week 5

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Day/Time | 9-10 | 10-11 | 11-12 | 13-14 | 14-15 | 15-16 | 16-17 |
| Monday | 1. Reporting & visualising (Academic Skills) 2. Interactive Workshop 3. Interactive Mock Assessment | | | | | | Q&A and day recap |
| Tuesday | DataLab 04: Conclusion  Summarize the problems statement, research questions, what the data implications were for these and how to proceed. Potential references. | | | | | | Day reflection |
| Wednesday | 1. Introduction to regression & machine learning (supervised, unsup..., the CRISP DM model) 2. Interactive Workshop (short) 3. Interactive Mock Assessment (short) 4. Working on creative brief poster | | | | | | Q&A and day recap |
| Thursday | 1. Data Science Block A recap 2. Working on creative brief poster 3. Deadline at 17:00 | | | | | | Q&A and day recap |
| Friday | DataLab 05:  Poster presentation day  Peer reviewed Poster Assessment. | | | | | | Day reflection |

## Assignment: Poster-Our world in data

In order to change the world, we must first be able to measure it. Measuring and quantifying what matters is a key component of evaluating progress. In September 2015, all countries of the world agreed to adopt certain goals as targets or indicators for global development. Collectively, these goals are known as the United Nations Sustainable Development Goals (SDGs). In this assignment, you are tasked with evaluating and describing [**our world in data**](https://ourworldindata.org/), in particular investigating how the world is performing on its sustainable development goals.

Graphical user interface, application

Description automatically generated

Figure 2: The United Nations Sustainable Development Goals.

Students are expected to adhere to a data science lifecycle process flow (see Figure 3).

The stages of the project (which will be distributed across the data lab days) are:

* **Define a data-driven research question**

Students are expected to frame a data-driven related research question related to a sustainable development goal. A question which can be answered on the basis of data-driven arguments is a data-driven research questions. Questions can range from simple (e.g., how is country X performing on indicator Y) to complex (e.g., how is country X performing on indicator Y and what are they key variables related to Y). Use [SDGTracker](https://sdg-tracker.org/) to help formulate your problem statement.

* **Gather the appropriate dataset**

Students are expected to use the sustainable development goals data bank (see Figure 4) in order to gather the appropriate dataset to answer their research question. Use [SDGDataBank](https://ourworldindata.org/sdg-tracker-update) to gather the appropriate dataset.

* **Explore your dataset**

Students are expected to apply the exploratory data analysis (EDA) concepts learned in the workshops to their datasets.

* **Analyze your data and present your findings using visuals**

Students are expected to apply concepts learned in the workshops and generate appropriate visuals which help support their arguments towards answering their research question.

* **Provide a solution to your research question with supporting arguments (based on your analysis)**

Students are expected to answer their research question with supporting arguments based on the work conducted in the previous data labs. Further, students are expected to summarize their analysis in the form a poster.

* **Poster Presentation Day**

Students are expected to present their poster to staff and peers. Students and staff will be allowed to vote for the best poster.

Figure 3: Project Lifecycle

## Challenges

## Requirements

|  |  |  |
| --- | --- | --- |
| **Genre** | Conference Poster Presentation | |
| **Brief description** | As a data scientist you will expected to: 1) Translate a use-case (e.g., a business requirement) into a data-driven research question. 2) Present your solution to multiple stakeholders.  Posters are an informal alternative to a full presentation. In addition, creating a poster forces you to think more about including information that is relevant, while discarding information that may not be relevant. Finally, posters are an ideal means to unleash your creative side. A badly designed poster, even with a sophisticated data analysis, remains a badly designed poster. | |
| **Software** | Use R to read, explore and visualize the data. Use a software of your choice to create the poster. | |
| **Poster  Components** | 00 Introduction | What is the problem statement? |
| 01 EDA Methodology | Describe the data. |
| 02 Findings | Include relevant visuals. |
| 03 Discussion | Do we have a data-driven solution to the problem statement? |
| 04 Conclusion | Critically look back at your solution. |
| 05 Presentation and additional information. | Sources, references, notes and author information. Live presentation of poster on DataLab 05. |
| **References** | [Making a better research poster - YouTube](https://www.youtube.com/watch?v=AwMFhyH7_5g) | |

For a more detailed explanation of expected contents of the posters; and an example of a data science poster template:

[**DS1-AssesmentPoster\_Template.pptx**](https://edubuas.sharepoint.com/sites/DevTeamAAIDM/_layouts/15/Doc.aspx?OR=teams&action=edit&sourcedoc=%7b217F0ADE-5F00-4ADB-B45E-CF4F75175AE5%7d)

## Deliverable

The posters are to be handed in Microsoft Teams or Github no later than 5pm on the day before the session in which we present the posters. There is no specified file format or size for handing in your poster but confer with a lecturer beforehand if you're handing in something other than PowerPoint, PDF or word to make sure they can actually open the file.

Our world in data - SDG Indicators


Figure 4: SDGTracker Data Bank.

# Creative Brief III – Image Data processing using python

## Timeline

### Week 6

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Day/Time | 9-10 | 10-11 | 11-12 | 13-14 | 14-15 | 15-16 | 16-17 |
| Monday | Programming, Algorithms, Flowchart | | | IDE setup, Python setup, variables, conditional statements, functions | | | Q&A and day recap |
| Tuesday | Foundation exercises with Python using Functions | | | | | | Day reflection |
| Wednesday | Data types, operators | | | Data structures | | | Q&A and day recap |
| Thursday | File operations / libs and Panda | | | Strings, texts, csv, json | | | Q&A and day recap |
| Friday | Basic file operations & string operations with files | | | | | | Day reflection |

### Week 7

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Day/Time | 9-10 | 10-11 | 11-12 | 13-14 | 14-15 | 15-16 | 16-17 |
| Monday | Data structures - II | | | NumPy | | | Q&A and day recap |
| Tuesday | Basic Image data processing - e.g.*: color modifications and bit masking* | | | | | | Day reflection |
| Wednesday | Bit/Logical operators | | | Vectorisation and broadcasting | | | Q&A and day recap |
| Thursday | Matplot – lines, points | | | Scikit-learn | | | Q&A and day recap |
| Friday | Image Processing using standard python libraries | | | | | | Day reflection |

### Week 8

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Day/Time | 9-10 | 10-11 | 11-12 | | 13-14 | 14-15 | 15-16 | 16-17 |
| Monday | Assignments to Python foundation | | | | | | | Q&A and day recap |
| Tuesday | Image Steganography | | | | | | | Day reflection |
| Wednesday | Assignments to Python foundation | | | | | | | Q&A and day recap |
| Thursday | Assignments to Python foundation | | | Application development – Flask, Qt5 | | | | Q&A and day recap |
| Friday | Image Steganography | | | | | | | Day reflection |

## Assignment: Image Steganography using Python

Steganography is a technique by which information is concealed in plain sight. There are many ways to conceal information. One of the most common techniques is to embed information (text, images, sound) in images. A commonly used method for image steganography the least significant bit (LSB) method. One of the key advantages of encoding information in the least significant bit is that the human eye cannot distinguish between the original image and the image with secret information embedded in it.

In this assignment, you will develop a steganography algorithm which can conceal secret information in an image.



Figure 5: Embedding information in the Mona Lisa

## Challenges

## Requirements

* Only use standard python libraries such as Pandas, Scikit\_learn, NumPy and Matplot-lib.
* Your code must consist of an encoder function (to embed the secret message) and a decoder function (to decode the secret message).
* Code must be well formatted and commented. Use the PEP8 style guide.
* At time of submission, code should compile with no errors.
* Use Github for version control.
* Use Jupyter notebooks.

## Deliverable

The jupyter notebooks are to be uploaded to Github no later than 5pm on last data lab day. Confer with a lecturer beforehand if you're handing in something other than jupyter notebook.

## References

K. Thangadurai and G. Sudha Devi, "An analysis of LSB based image steganography techniques," 2014 International Conference on Computer Communication and Informatics, 2014, pp. 1-4, doi: 10.1109/ICCCI.2014.6921751.