



## Data Synthesis Course Descriptor

Course Title	Data Synthesis	Faculty	EDGE Innovation Unit (London)
Course code	NCHNAP565	Course Leader	Professor Scott Wildman (interim)
Credit points	15	Teaching Period	This course will typically be delivered over an intensive 2-week period.
FHEQ level	5	Date approved	June 2020
Compulsory/ Optional	Compulsory		
Pre-requisites	None		
Co-requisites	None		

### COURSE SUMMARY

This course is an intensive two-week, face-to-face bootcamp, which covers how data is brought together from different sources (text, audio, images etc) to enrich data science analysis. The course focuses on two aspects: parsing images using Computer Vision and parsing language, using Natural Language Processing (NLP). Learners will gain hands-on experience in using Computer Vision and NLP and will explore how to extract information from a variety of formats to inform data analysis.

### COURSE AIMS

- Train learners in how to use data warehousing for effective data science solutions.
- Train learners in how to synthesis data using a wide variety of sources and formats, including audio, text and images.
- Introduce learners to computer vision and natural language processing.
- To learn from other data scientists and become part of a wider data science community.

## LEARNING OUTCOMES

On successful completion of the course, learners will be able to:

### KNOWLEDGE AND UNDERSTANDING

- K1b Understand the key principles, concepts and syntax of NLP and Computer Vision.
- K2b Evaluate and critique technical tools and processes for combining diverse sets of data.

### SUBJECT SPECIFIC SKILLS

- S1b Use NLP and Computer Vision to parse language and images to solve complex data science problems.
- S2b Effectively use data synthesis to bring diverse and complex datasets together for analysis.
- S3b Work collaboratively to complete complex data synthesis tasks.

### TRANSFERABLE AND PROFESSIONAL SKILLS

- T1b Create visual presentations to a high standard.
- T2b Manipulate, structure and transform complex data.
- T3bi Critically investigate, make judgements and use self-initiative.
- T3bii Demonstrate an effective technical proficiency of written English that uses a wide range of literacy skills and vocabulary selected appropriately to communicate to specialist and non-specialist audiences.

## TEACHING AND LEARNING

This is a face-to-face bootcamp, of two weeks duration, taught once every year.

This course can be offered as a standalone short course.

Teaching and learning strategies for this course will include:

- Lectures
- Informal discussion groups
- Practical sessions
- Assessment

Course information and supplementary materials will be available on the College's Virtual Learning Environment (VLE).

Learners are required to attend and participate in all the formal and timetabled sessions for this course. Learners are also expected to manage their self-directed learning and independent study in support of the course.

The course learning and teaching hours will be structured as follows:

- Off-the-job learning and teaching (12 days x 7 hours) = 84 hours
- On-the-job learning (10 days x 7 hours) = 70 hours

Apprentices will complete workplace activities before and after the bootcamp, as part of their on-the-job learning. Preliminary activities include reading and workplace research and post-bootcamp activities will include completion of the assignments (see below).

## ASSESSMENT

### FORMATIVE

Learners will be formatively assessed during the course by means of set assignments. These will not count towards the final degree but will provide learners with developmental feedback.

### SUMMATIVE

Assessment will be in two forms:

AE	Assessment Type	Weighting	Online submission	Duration	Length
1	Project	70%	Yes	Requiring on average 25-35 hours to complete	-
2	Presentation	30%	Yes	30 mins	-

### FEEDBACK

Learners will receive formal feedback in a variety of ways: written (via email or VLE correspondence); oral and indirectly through online discussion groups. Learners will also attend a formal meeting with their Academic Mentor (and for apprentices, including their Line Manager). These bi- or tri-partite reviews will monitor and evaluate the learner's progress.

Feedback is provided on summatively assessed assignments and through generic internal examiners' reports, both of which are posted on the VLE.

## INDICATIVE READING

Note: Comprehensive and current reading lists for courses are produced annually in the Course Syllabus or other documentation provided to learners; the indicative reading list provided below is used as part of the approval/modification process only.

### BOOKS

- Forsyth, D., (2012), *Computer Vision: A Modern Approach*, Boston, Mass.; London: Pearson; Harlow: Pearson Education
- Eisenstein, J., (2019), *Introduction to Natural Language Processing*, Cambridge, Massachusetts: The MIT Press
- Farzindar, A., (2020), *Natural language processing for social media*, San Rafael, California: Morgan & Claypool

### JOURNALS

Learners are encouraged to consult relevant journals on data synthesis, Computer Vision and/or NLP.

### ELECTRONIC RESOURCES

Learners are encouraged to consult relevant electronic resources on data synthesis, Computer Vision and/or NLP.

**INDICATIVE TOPIC**

- Computer Vision for data science
- NLP for data science
- Combining diverse data and formats

<b>Title: NCHNAP565 Data Synthesis</b>					
<b>Approved by: Academic Board</b>					
<b>Location: Academic Handbook/Programme specifications and Handbooks/ Undergraduate Apprenticeship Programmes/BSc (Hons) Data Science Programme Specification/Course Descriptors</b>					
Version number	Date approved	Date published	Owner	Proposed next review date	Modification (As per AQF4) & category number
2.1	May 2022	May 2022	Scott Wildman	September 2026	Category 1: Corrections/clarifications to documents which do not change approved content.
2.0	January 2022	April 2022	Scott Wildman	September 2025	Category 3: Changes to Learning Outcomes.
1.0	June 2020	June 2020	Scott Wildman	June 2025	