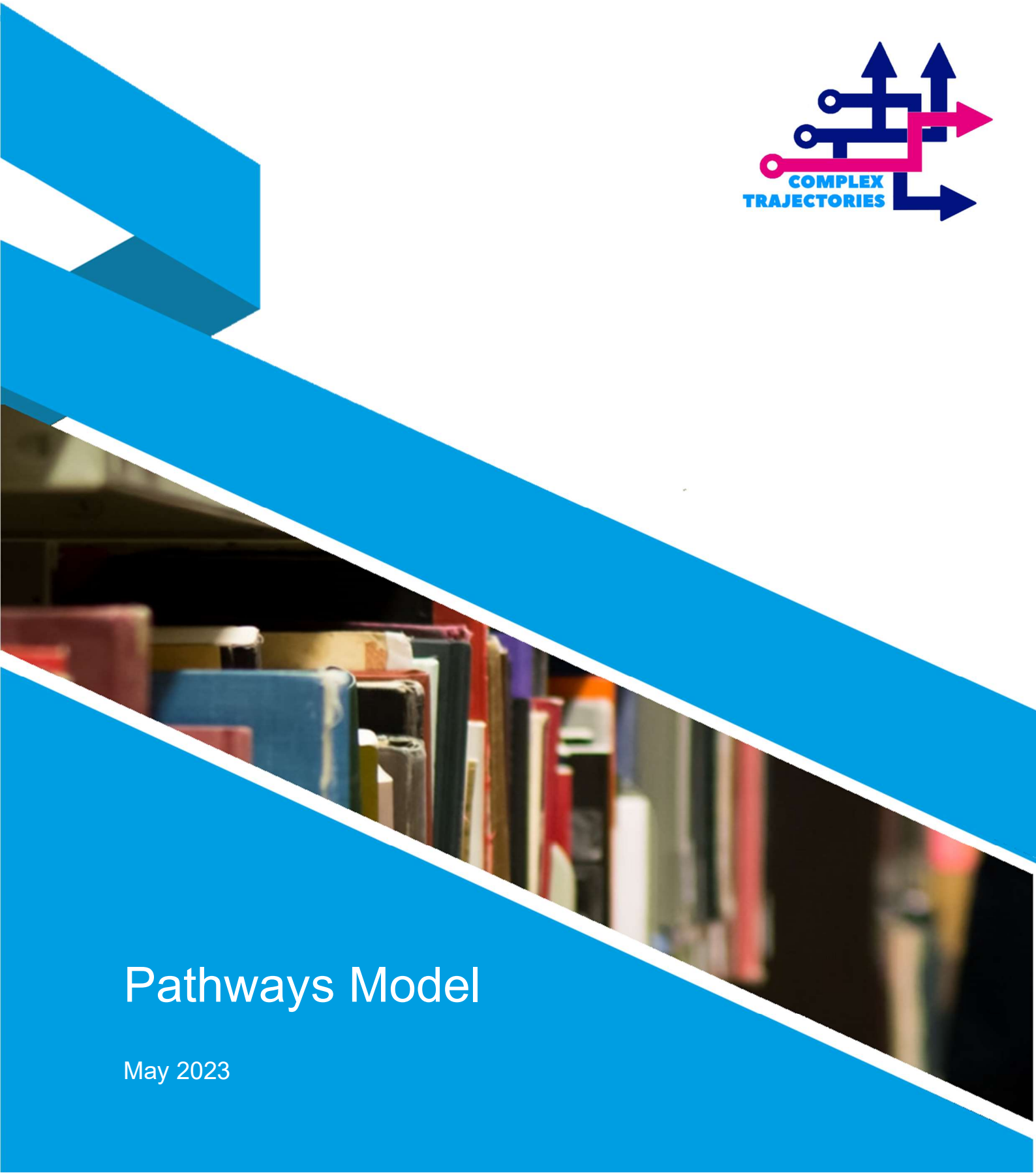


Pathways Model

May 2023

Co-funded by the
Erasmus+ Programme
of the European Union

Author:

Chris Edwards

The following documents are embedded in a whole MOOC course on Longitudinal analysis techniques developed in the framework of the same Complex Trajectories project. Those interested in doing the MOOC should consider the following procedure:

1. Access the AULAbERTA space through the link: https://aulaberta.uab.pt/

2. Register in the platform a. Select one of the available languages (Portuguese or English) b. Follow the instructions given in the platform

3. After creating the account, they should access the MOOC Longitudinal Analysis through the link: https://aulaberta.uab.pt/course/view.php?id=94

Co-funded by the Erasmus+ Programme of the European Union

The Pathways Model: an introduction

Chris Edwards, Institute of Educational Technology, The Open University, UK

Co-funded by the Erasmus+ Programme of the European Union

Hello.
This presentation is a brief introduction to the Pathways Model developed at the Open University in the UK.

Contents

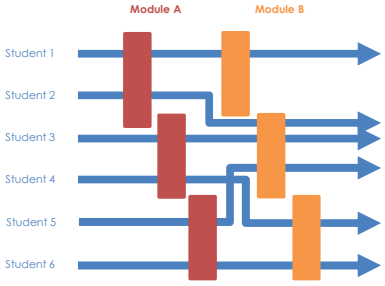
Co-funded by the Erasmus+ Programme of the European Union

I will answer three questions:

1. Why develop the Pathways Model?
2. What is the Pathways Model?
3. What are the benefits?

Why develop the Pathways Model? ...(i)

- Universities that grow to have **large** numbers of students can offer **larger** curricula
- The **increased** prospectus offers the potential for;
  greater choice **of** qualifications
  greater choice **within** qualifications
- Universities may also offer greater flexibility to move **between** subjects – even an '**open**' degree
- the pathways picture becomes very **complex**



Co-funded by the Erasmus+ Programme of the European Union

The first question is why we developed the Pathways model.

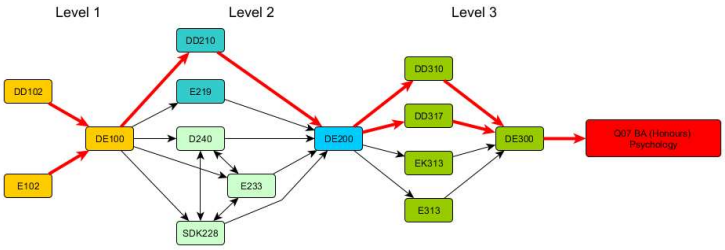
Over time, universities can grow to have **large** numbers of students which can support a large curriculum offer.

A large prospectus has the potential to offer students a greater choice **of** qualifications and a greater choice **within** qualifications.

Alongside this, a university may also decide to offer greater flexibility to move **between** subjects – perhaps even offering an **open** degree with very high levels of choice the pathways picture becomes very **complex.**

Why develop the Pathways Model? ...(ii)

- Example: qualification with relatively **simple** structure available in 2019
  This has **64** alternative study routes designed in
- Students can choose to study at **different** rates
  Greatly increases **actual** number of different study paths experienced



Co-funded by the Erasmus+ Programme of the European Union

An example of a version of a qualification with a relatively simple structure is the Psychology degree from the Open University's 2019 prospectus.

With just one element of choice at each level. That is during each year of a full time degree. We find that there are **64** alternative study routes designed in.

Students can choose to study at **different** rates. This greatly increases **actual** number of different study paths experienced.

Why develop the Pathways Model? ...(iii)

- This choice increases **personalisation** of study paths for students leads to **fragmentation** of cohort; large numbers of **different experiences**
- Measuring the **effectiveness** of curriculum requires an understanding of the **complexity of experience**
- Pathway questions can be answered with **current** data structures but may require **high** levels of resource
- What data structures would we use if setting up **now**? with different **requirements**, different **technologies**

Co-funded by the Erasmus+ Programme of the European Union

We can summarise this by saying that an increased level of choice increases the **personalisation** of study paths for students but that this leads to **fragmentation** of cohort and to large numbers of **different experiences.**

An institutional impact of this is that any attempt to measure the **effectiveness** of curriculum requires an understanding of the **complexity of experience.**

We have found, from experience, that whilst pathway questions can be answered with **current** data structures, **high** levels of resource are required.

And so, we asked the question, 'What data structures would we use if setting up **now:** with our different **requirements**, and with different **technologies** available?'

What is the Pathways Model? ...(i)

- Take the **student** perspective,
  put their experience **central** to the data structure
- Recognise that credit is linked to individual **modules**
  a student gains credit on **passing** a module
- Recognise that students have a choice of **presentation** for each module
  cohorts **fragment** over time
- These combine to become the **Smallest Creditable Element**

Co-funded by the Erasmus+ Programme of the European Union

Let us now take an overview of the Pathways Model.

Firstly, we tried to take the **student** perspective of their study experience putting this **central** to the data structure.

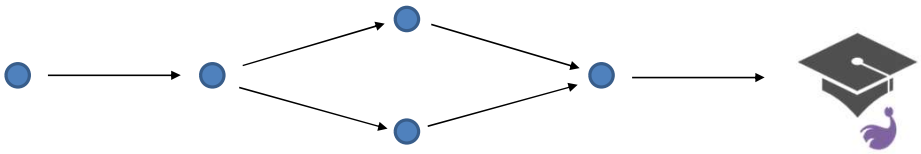
Going back to first principles, we recognise that credit is linked to individual **modules** and that a student gains credit on **passing** a module but that other outcomes are also possible.

We also recognise that students have a choice of **presentation** for each module and when they exercise this freedom, cohorts inevitably **fragment** over time.

These three factors can combine to become the **Smallest Creditable Element.**

What is the Pathways Model? ...(ii)

- Each value of the **Smallest Creditable Element** (Student + Module + Presentation) is **unique**
  a student can only have one **attempt** at each module presentation and there can only be one **final outcome**
- Students aim to pass successive module attempts
  until they achieve their qualification (or other learning goal)



Co-funded by the Erasmus+ Programme of the European Union

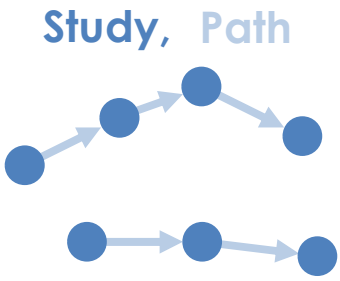
Each value of the **Smallest Creditable Element** is **unique. A** student can only have one **attempt** at each module presentation and there can only be one **final outcome.**

Students aim to pass successive module attempts until they achieve their qualification (or other learning goal).

What is the Pathways Model? ...(iii)

Data structure

- Using **graph** database technology
- Model as metaphor



Co-funded by the Erasmus+ Programme of the European Union

We chose to use graph database technology because of the way it supports relationships within data and the readily available visualisation options.

A graph database has two types of data structure:

1) firstly, **nodes** of information and
2) secondly, **edges** (with their own information) connecting the nodes.

The Pathways Model centres on two collections:
One is called **Study**. This is a collection of **nodes** and is essentially a table of the details for each value of the Smallest Creditable Element.
The other is called **Path**. This is an **edge** collection, that describes all the links between the nodes in Study.

This data structure is a metaphor for the student study experience.

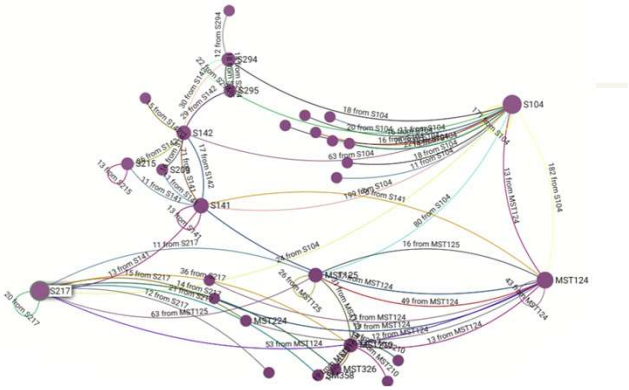
When we run a query the database presents the results **graphically**, like this example which is a study path for one student.
Each **node** has **all** the data for that **study attempt.**
Each **edge** has **all** the data for that **study gap.**

A simple toggle will move between presenting the visualisation and the data.

What is the Pathways Model? ...(iii)

Queries that aggregate data

- Example: find the ongoing study paths for the students starting with module A in presentation B
- The database recognises the connections and natively plots the results
  each **node** and **edge** contain all the aggregated data and can be explored further



Co-funded by the Erasmus+ Programme of the European Union

The real power of the model is realised through queries that aggregate data.

For example. This plot is the output of a summary of all the ongoing study paths for the cohort of students starting their undergraduate study with, let's say, module A, in presentation B.

The database recognises the connections within the data and natively plots the results.

Each **node** and **edge** contains all the aggregated data and can be explored further.

What are the benefits?

- The model as metaphor makes it **easier** to discuss with colleagues who are unfamiliar with databases and structures.
  The conversation is more **natural**
- The graph database approach embeds the relationship between module study attempts
  greatly **facilitates** pathway analysis
- Reveals the **complexity** within the student experience data
  offering a **new perspective** and
  a route to identifying the key questions for **improving** this experience

Co-funded by the Erasmus+ Programme of the European Union

Finally, we outline some of the benefits of using the Pathways Model:

1) The model as metaphor makes it **easier** to discuss with colleagues who are unfamiliar with databases and structures.
The conversation is more **natural** than one involving traditional databases.

2) The graph database approach embeds the relationship between module study attempts greatly **facilitating** pathway analysis.

3) The Pathways Model reveals the **complexity** within the student experience data offering a **new perspective** and a route to identifying the key questions for **improving** this experience.

Thank you

Please refer to the supporting materials on the MOOC website

11

Thank you .

Please refer to the supporting and all the other materials on the MOOC website.

The Pathways Model: in more detail

Chris Edwards, Institute of Educational Technology, The Open University, UK

Co-funded by the Erasmus+ Programme of the European Union

Hello, this video follows on from the Introduction to the Pathways Model and aims to give more detail about the way in which the model is set up and used.

Contents

Co-funded by the Erasmus+ Programme of the European Union

After a reminder of what was covered in the previous video, we look at our **choice** of database technology;

The **two** central collections, and then move into the database to look at these more closely and the way in which **queries** are entered and used, including how results are displayed;

Finally, we highlight the need for a simple notation to describe a student trajectory through a series of modules and our solution.

Recap from Introduction video

- Why the Pathways Model
  - Complexity arises from increased student choice in an age of personalisation
    We need to facilate an **understanding** of student experience
    new tools to **evaluate** effectiveness of curriculum
- Key elements of Pathways Model
  - Smallest Creditable Element (SCE)
  - model as metaphor (graph technology)
- Benefits of Pathways Model
  - **facilitates** discussion with colleagues and students
  - **powerful** queries
  - opens window to **understand** complexity of student experience

Co-funded by the Erasmus+ Programme of the European Union

In the previous video we explained that the Pathways Model was needed to facilitate our **understanding** of the student experience in an age of increased personalisation and provide new tools for evaluating the effectiveness of curriculum within this complexity.

We described the key elements of the **Smallest Creditable Element** and the idea of the model as **metaphor,** and outlined the **main** benefits of the model. That it:

- **facilitates** discussion with colleagues and students,
- provides **powerful** queries, and
- opens window to **understand** the complexity of student experience.

The database technology

- We have chosen to use ArangoDB
  - multimodel - graph, document, and key-value
  - completely flexible – structure not fixed
  - single SQL-like query language
  - range of options from:
    - free community edition running on single computer
    - to secure cloud-based managed institutional solution
  - continued development
  - Open Source
  - JSON data format

Co-funded by the Erasmus+ Programme of the European Union

There are several potential databases available. We have chosen to use **ArangoDB.**

This is a **multimodel** database combining - graph, document, and key-value structures.

It is completely **flexible** – the structure is not fixed and can be modified at any time to match changing needs.

It has a **single** SQL-like query language.

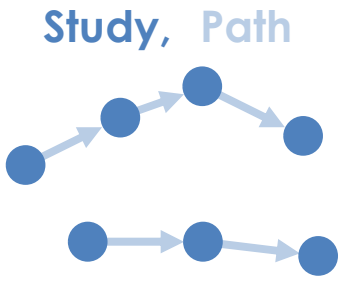
It is available in a **range** of options from:

- the **free** community edition that can run on single computer
- to a secure cloud-based managed **institutional** solution

It is under **continued** development, is Open Source and uses JSON data format.

The two collections

- Study:
  - document collection
  - each document contains:
    - unique value of Smallest Creditable Element
    - all data relating to the study attempt
- Path:
  - edge collection
  - each document contains:
    - study attempt at start, and at end of edge
    - other data relating to study gap



Co-funded by the Erasmus+ Programme of the European Union

In the previous video we explained that the Pathways Model centred on two data collections:

The first is **Study** which is a **document** collection. Each document contains the **unique** value of that **Smallest Creditable Element** and **all** data relating to that student's **study attempt.**

The second is Path. This is an **edge** collection. Each document lists the study attempt at the **start**, and the one at the **end** of the edge. Plus, **other** data relating to that **study gap**

Let's move into the database to view these collections and explore queries and their output...

Demonstration in ArangoDB

- This section of the presentation is a screen recording of using ArangoDB in its browser interface. It shows:
  - the Collections view, including the
    - **Study document** collection
    - **Path edge** collection
  - The Query view, and how the **Traversal** query works

Co-funded by the Erasmus+ Programme of the European Union

In ArangoDB browser interface

Let us take a look at the browser interface for ArangoDB. Down the left-hand side there is a menu and the first item is **Collections**, which is currently selected. In the main panel we see a button to add a collection within a tabular list of all the collections. There are two icons. The collections with the document icon are document collections. These are similar to tables in a conventional database. The collections with the network symbol are edge collections. These contain the sets of edges.

Let's look at the **DemoStudy** document collection. We see a menu appears along the top of the panel. It defaults at Content. Below this, the content is arranged in rows. We can go to the end. Let's look at one document. The first thing we notice is a header with green items. These uniquely identify the documents. Note there are three entries. Below this there's a menu. And below this we see there's a list of 76 objects. Each object is a combination of a name and a value. And as we go down, we can work through them. When we get to this one, 'cont', which is short for continuous assessment, we can expand it by clicking on the triangle. And then expand each assessment to find the cut off dates and the scores.

Now, we go to the DemoPath collection. Here, **DemoPaths2**. This looks very similar. And if we open one of the records, we find there are now five green items at the top. In addition to the three for uniquely identifying the documents, we find there's a 'from' and a 'to'. These detail the nodes at the start of the edge and at the end of the edge. Both of these refer to vertices within the DemoStudy collection.

The next most important tab is Queries. Here we have probably the most important query that we'll use. It's the **Traversal** query, and if we read it. It says,

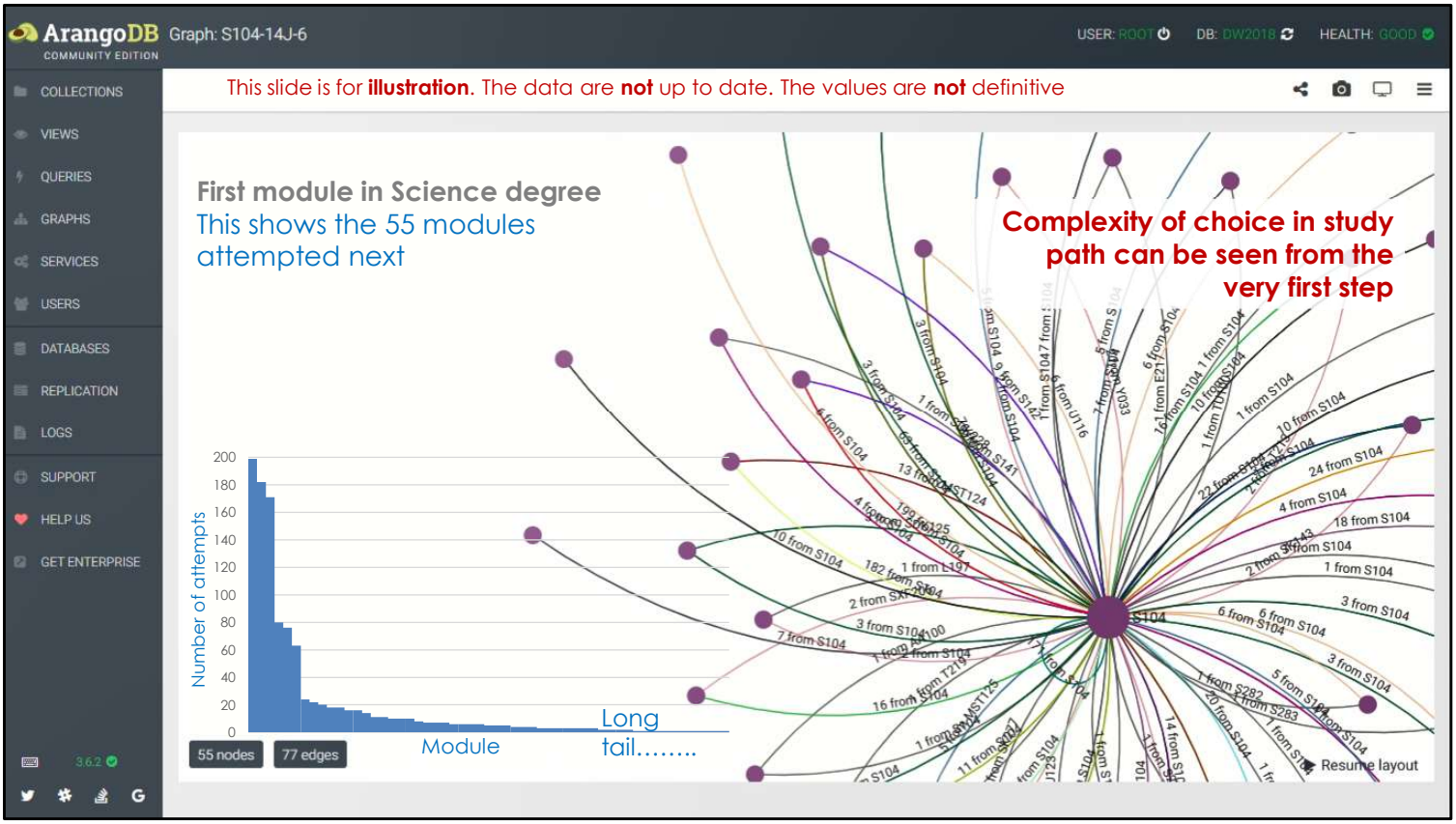
> *'For all the vertices, edges and paths, between 1 and 10 steps outbound from a starting node, using the edges in DemoPath2 return (in this case we have) 'p' for path.'*

If we change that to 'v' for vertices and add this first row that defines the starting node. We can run it. We see that the results are returned in the panel below. These are all the vertices that this student has studied. There are 44 which is more than the student has actually studied, and that's because some are returned more than once. We can simplify this. Collecting all the vertices that are equal together and then returning those. Running it again and we now find we have the 11 nodes representing the actual modules that this student has studied.

We can do the same for edges. So if we change the vs to es, and run it again, we now find there's fourteen different edges involved. That's because of this here. Hang on, let's just spread this out a little bit. There we go, and we see there's a direction from the start and the arrow shows where this path goes. We see this arrow has double heads. It's a double ended arrow. Actually two arrows, one over the other. One going one way and one going the other. These two modules were studied together by the student, concurrently, at the same time.

The database recognised that all these are related and naturally presented them as a as a graph. We can click on the JSON tab here or toggle here, and and switch it to the JSON output. So, this is a list of all the edges. Let's just clear all those and now we can move to paths. This Collect function doesn't reduce the number completely, so we'll leave it out. For this we need to use a different filter. When we're looking at paths. And when we look at the paths, we've gone back to 44. So, some are repeated, but we won't worry about that, just at the moment.

We recognise this pattern because it's showing the study for this particular student and we can hover again and get each individual node. And we can click on JSON and here we find that the elements are arranged in groups. So, we have the group of vertices to start with and then edges, and then we can also introduce weightings to each of these edges, and so we have weightings, there.



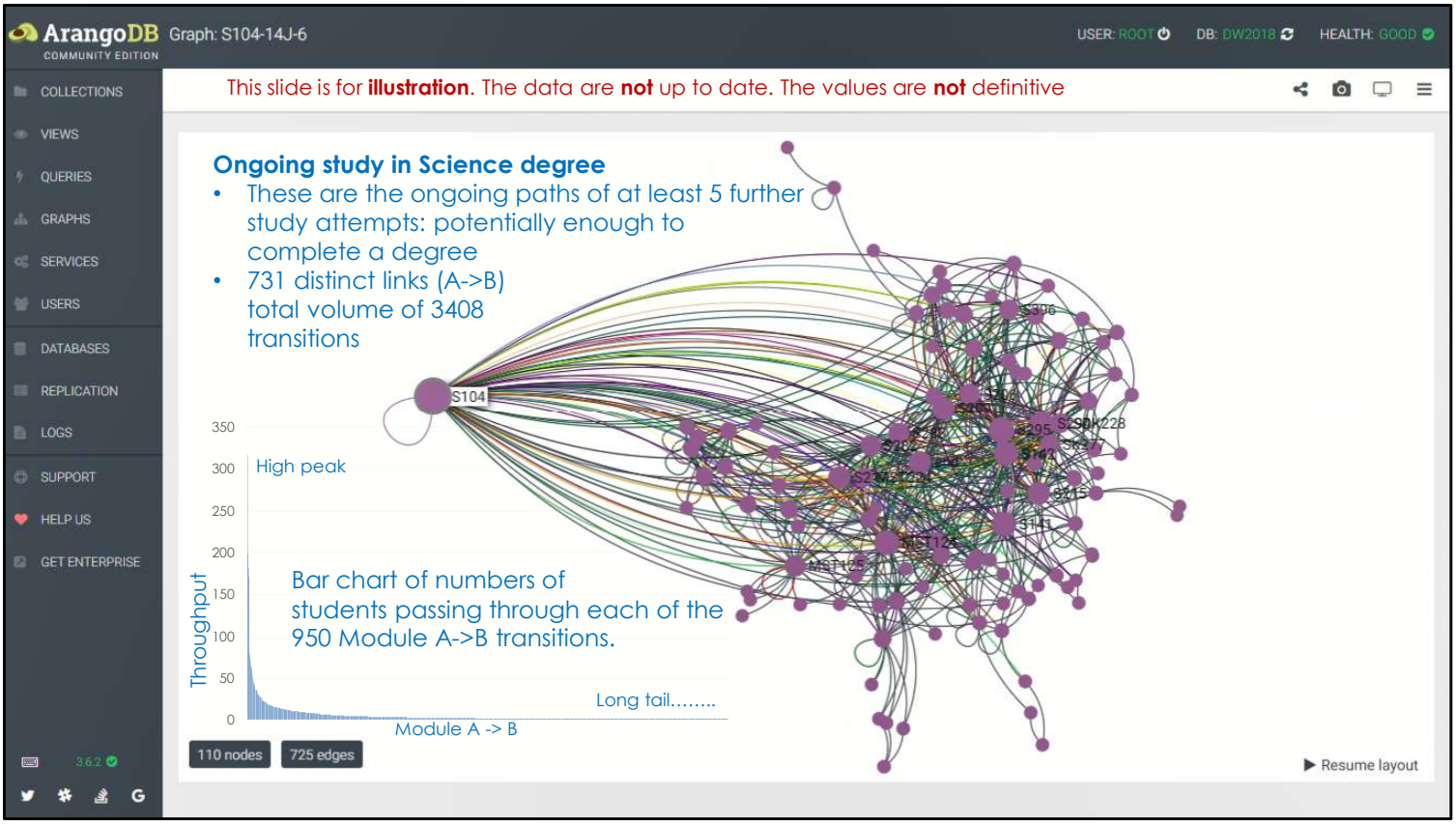
Here, we demonstrate the output of a **graph** query and have overlayed some text.

The graph shows the next module studied after a cohort of students began their **first** module **together** on a Science degree.

The bar chart shows, arranged in descending order, the number of students **attempting** each of the 55 modules.

We have found the shape of this plot is **typical** of any analysis of study path.

It has the **majority** of students following a **small** number of paths alongside a **significant** proportion of students overall, on much **less** popular trajectories.



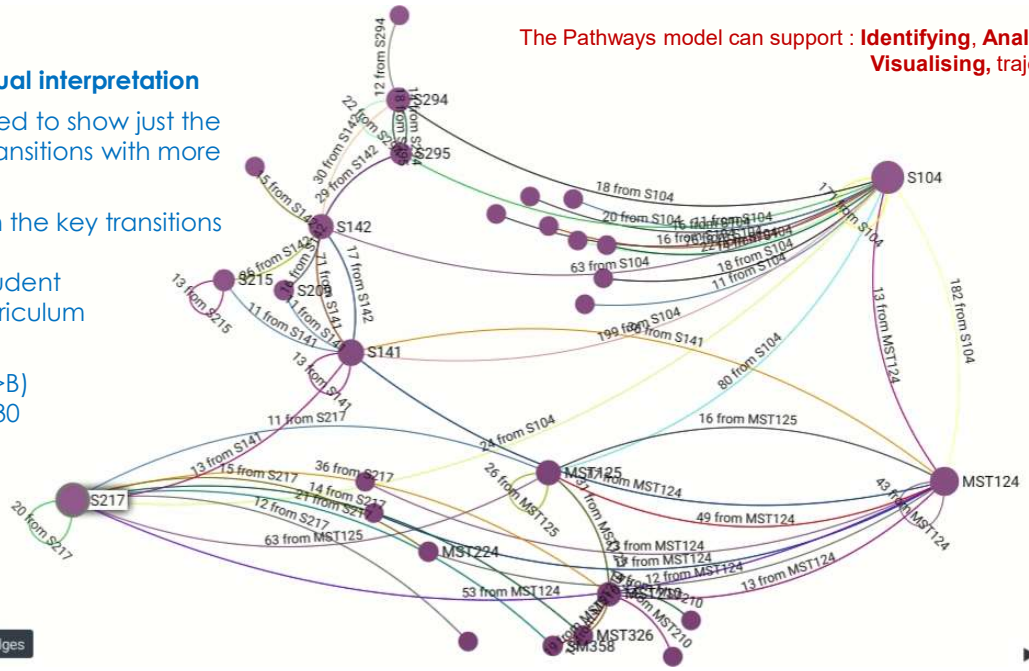
This plot shows the study trajectories generated by this same cohort of students as they study all the modules they need for a degree.
This clearly highlights the complexity and brings all the data together in a single view.
However, the plot is too dense to be of much use beyond this.
The bar chart shows the same pattern of numbers of students on each trajectory.

The student study path – busiest routes

The Pathways model can support : **Identifying**, **Analysing,** and **Visualising,** trajectory data

Filtering to enable visual interpretation

- As before but filtered to show just the module-module transitions with more than **10** students
- Enables a focus on the key transitions and modules that most impact on student outcomes and curriculum effectiveness
- 59 distinct links (A->B) total volume of 1930 transitions



Co-funded by the Erasmus+ Programme of the European Union

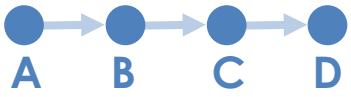
29 nodes 57 edges

▶ Resume layout

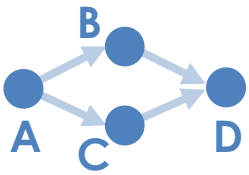
The ability to filter enables the visualisation to become a useful tool.
Here we select all the transitions that have been taken by at least 10 students.

Notation for pathways

- Graphical visualisations are very informative
- In order for analysis, discussion, and reporting
  Need a simple of way of describing a pathway in text.



A -> B -> C -> D



A -> (B,C) -> D

Co-funded by the Erasmus+ Programme of the European Union

Interactive graphical visualisations are very informative. However, in order to continue analysis, engage in discussion, and report on student trajectories, a notation is needed.

These two examples show the basics of the notation we have developed.

1. The first is to describe a series of four modules taken sequentially.
2. The second is for the same modules with B and C studied together, concurrently.

Thank you

Please refer to the supporting materials on the MOOC website

11

Thank you. I hope you have found this more detailed video interesting and useful.

Please refer to the supporting and all the other materials on the MOOC website.

The Pathways Model: examples of use

Chris Edwards, Institute of Educational Technology, The Open University, UK

Co-funded by the Erasmus+ Programme of the European Union

Hello.

This video is the third in a series of three. It follows on from the two previous presentations on the Pathways Model.

In it we will describe several examples of how the model has been used.

Contents

Co-funded by the Erasmus+ Programme of the European Union

We will look at three examples where the Pathways Model has offered insights that had not been revealed through other methods.

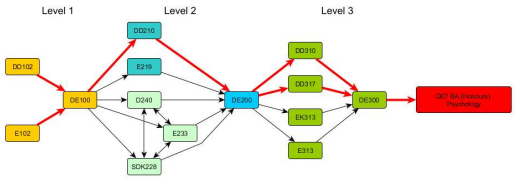
Firstly, an insight into the trajectories of Psychology students in a qualification review.

Secondly, revealing the complexity within the Language Studies degree.

Lastly, an exploration into the relationship between age and degree completion.

Exploration of a Psychology degree

- New Psychology degree was reviewed in 2018
  - introduced in 2012
  - all modules performing well
  - final module studied by over 2000 students



- Existing approaches did **not** consider student trajectories.
  Pathways Model **readily** able to trace all students paths through programme.
- **Not** one student, at that point had followed the programme of study all the way from beginning to end.
  The students **completing** had either transferred from previous qualificaiton or used credit transfer.

Co-funded by the Erasmus+ Programme of the European Union

When we introduced the Pathways Model in the first video, we also introduced the OU's Psychology degree,
as it was set out in the universities prospectus in 2018.

This is one of the simplest degrees structurally and we repeat the schematic of the programme of study here.

When the qualification was reviewed in 2018 it was found that all modules were performing well for students.
And that the final module (the single green module on the right) had been studied by over 2000 students.

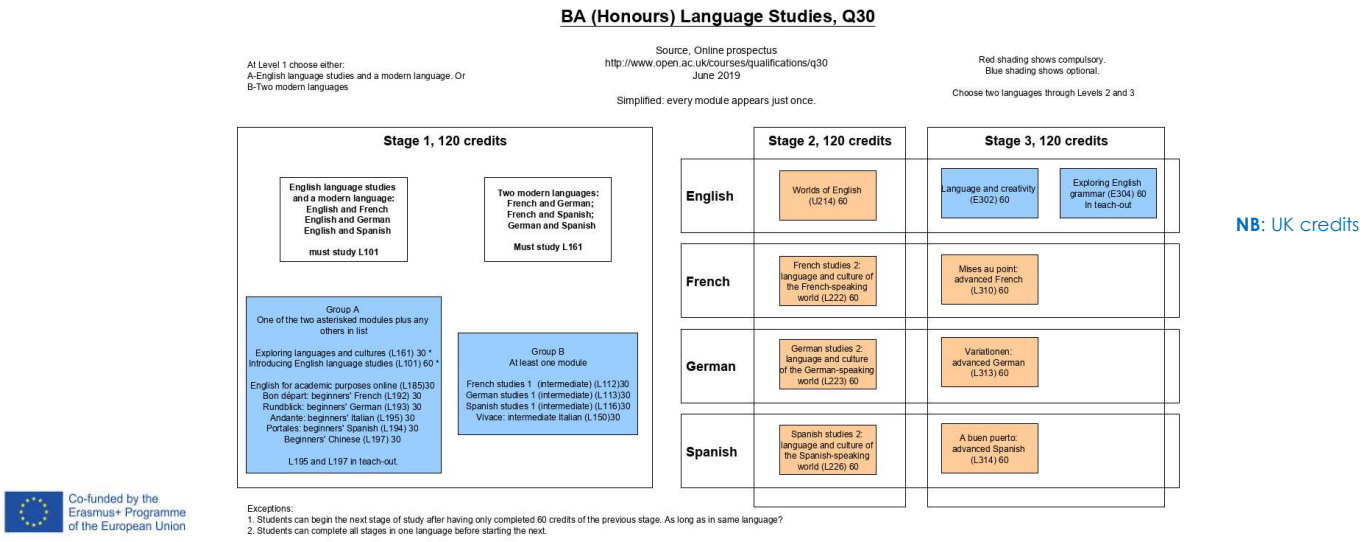
However, approaches existing at the time did **not** consider the study trajectories of students.

The Pathways Model was readily able to trace all the students following the programme of study.

At the time of the review, it was found that **no** student had yet followed the programme of study **all** the way through to the end.

Students completing this **young** qualification achieved it through counting modules from the previous version or from credit transfer.

Study trajectories in languages



NB: UK credits

Co-funded by the Erasmus+ Programme of the European Union

In contrast to the Psychology degree, the Language Studies degree was structurally significantly **more** complex in 2019.

Each stage of study corresponds to a year in a 3 year degree.

The options within Stage 1 design in **1436** potential study trajectories through this stage.

This contrasts with **just** 2 study routes through Stage 1 in Psychology.

Study trajectories in languages (ii)

- The options within Stage 1 design in **1436** potential study trajectories. In practice – **more**. Psychology had **just** 2 study routes through Stage 1
- At the time there was no tool available to review study paths.
  The Pathways Model:
  - revealed **236** students had commenced their degree with one of the Language modules and had made it as far as one of the Stage 3 modules.
  - These students had followed **224** different study trajectories.
    Four typical paths are shown. Each of these contains some **concurrent** study

(AA100-Pass) -> (L140-Pass, L192-Pass) -> (L204-Pass) -> (U214-Pass) -> (L314-Pass) -> (E302-Pass)
(AA100-Pass) -> (L140-Pass, L193-Pass) -> (L211-Pass) -> (L204-Pass) -> (L314-Pass) -> (L310-Pass)
(AA100-Pass, L120-Pass, L161-Pass) -> (L211-Pass, U214-Pass) -> (E302-Pass, L310-Pass)
(AA100-Pass, L140-Pass, L161-Pass) -> (L204-Pass, U214-Pass) -> (E304-Pass, L314-Pass)

Co-funded by the Erasmus+ Programme of the European Union

5

Although there were already many potential study trajectories through Stage 1 of Language Studies, in practice, there were many more. This is because of students' ability to choose which presentation of a module they study.

No one had reviewed student study paths because there had been **no** tool available and **no** practical way to do this.

The Pathways Model revealed that **236** students had commenced their degree with one of the Language modules and had made it as far as one of the Stage 3 modules.

These students had followed **224** different study trajectories.

Four typical paths are shown. Each of these contains some **concurrent** study.

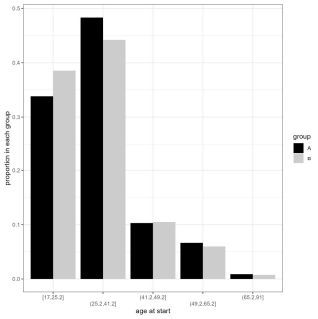
These are the modules **within** brackets.

The final row shows a student who had studied concurrent modules at **each** Stage.

This student had actually studied at a **higher** intensity than a normal **full-time** student.

A question of age...

- Are older students are **less** likely to achieve a degree than younger students?
- Pathways Model provides straightforward approach to approach this:
  - by identifying when students start their degree study and their **age at start**, and
  - those who **achieve** a degree
  - Comparing this age for both 'populations'
- There **is** an age-related factor that appears to operate within the **26-41** age group.



Co-funded by the Erasmus+ Programme of the European Union

The analysis for this slide provides a fresh example of how the Pathways Model is a valuable tool for any trajectory related question.

The question in this case is whether older students are less likely to achieve a degree than younger students.

Using the Pathways Model it is straightforward to build a query that identifies when students start their degree study, their age at start and those who achieve a degree.

Using this value of age to compare the two 'populations', we find that there **is** an age-related factor and that it appears to operate within the 26-41 age group.

Thank you

Please refer to the supporting materials on the MOOC website

7

Thank you.

I hope you have found this series of videos interesting, and that you find these examples both interesting and useful.

Please refer to the supporting materials on the MOOC website.