PROJECT GUIDE 10

Tree testing

Version: 1.4 (March 2010)

Also known as card-based classification evaluation, this is a quick, simple, and cost-effective technique to usability test a hierarchical classification (such as a site's information architecture). Included are guidelines on when to use this technique, how to prepare for it, how to analyse the results, along with the resources and time required, and the steps involved.

About tree testing

This is a quick, simple, and cost-effective technique used to usability test a hierarchical classification such as the information architecture (the structure) of an intranet.

During the activity, participants are provided with a set of short scenarios written on index cards. They are shown the items in the top level of a classification (also written on index cards) and asked to identify where they would look in the classification for each scenario. As they make a selection, a facilitator shows the next level of the classification and they again indicate where they would look. The participant 'drills down' through the hierarchy to identify the categories and sub-categories that allow them to complete the scenario.

It provides feedback on the usability of the classification, groupings and labelling, and provides insight into what should be included within categories. The outcomes from the activity are used to further refine a classification or to determine that it is ready to implement. .

Figure 1: Identifying a location in a hierarchical classification

When to use

Tree testing is a valuable technique to use when developing a classification such as the intranet's information architecture.

Use it when:

- a draft classification has been prepared and is ready for usability testing
- there are different ways of grouping information and you would like to find out which method works best for the user group
- there are a number of labels that could be used to describe information and you would like to check which is most easily understood
- content could be included in more than one place in the classification and you would like to determine which location is best

Important note

Tree testing is one method of involving users in an intranet design. It does not test all aspects of usability and should be combined with user centred design techniques and additional usability tests. See *Related techniques* (page 3).

Key strengths

Tree testing has the following strengths:

- It takes little time for the participants, who only need to give you ten to fifteen minutes to provide valuable input. This means that participants are simple to recruit and are happy to be involved.
- It focuses on real tasks that people would undertake. In contrast, activities such as card sorting focus only on the content, and the outcomes may not allow users to complete tasks easily.
- It is simple and easily understood. It can be explained to participants in a matter of minutes.
- It is quick and cheap to run. The evaluation takes few resources and a short amount of time so can be done as needed.
- It is low-tech and flexible. The classification can be modified during the evaluation if necessary to test alternative approaches.
- It needs no special training for organisers.
- It involves users in the design process, and demonstrates that the intranet will be created with their needs in mind.

Important note

Tree testing does not test all aspects of a classification. For example it does not test whether people can locate related information. It is also only suitable for testing simple tasks, not complex information tasks.

Related techniques

Tree testing is one input into a user-centred design process. It is best teamed with the results of other analysis, design and evaluation techniques, such as:

- Task analysis

  Task analysis focuses on the tasks that people want to undertake using the intranet and the type of information and activities that are part of the task. Task analysis is used as an input to the development of an information architecture and is a source for the scenarios to be used in tree testing.
- Card sorting

  Card sorting is a user-centred design technique used as input in the development of an information architecture.
- Usability testing

  Usability testing is another task-based evaluation technique that can be used to validate the information architecture and page layouts for the intranet

> **Important note**
>
> Card based classification evaluation should not be the only usability test for a site redesign as it does not test the page layouts, navigation and search function for the intranet. It only tests simple tasks and does not consider tasks where a range of information may be necessary.

Pre-requisites

Before using this technique, the following activities need to be conducted:

- Conduct needs analysis

  Needs analysis identifies the information needs of users. The outcomes from the needs analysis can be used to generate scenarios for the design of the information architecture and for usability testing activities.

  Needs analysis also identifies the staff groups that will use the intranet, and therefore the type and number of participants required for the evaluation.
- Develop scenarios

  Scenarios are short stories describing a user task involving the intranet. Scenarios are an integral part of any usability testing activity as all usability testing involves participants working through a number of scenarios.

  Scenarios are developed as an outcome of needs analysis activities.
- Develop draft classification

  A draft classification or information architecture is required. The information architecture will generally be a hierarchy (or number of hierarchies). The top 1–2 levels at least must be included in the draft, and any additional levels as needed. The information architecture can be relatively unfinished. There may be some content groups included in more than one place in the hierarchy (to test which works best) and alternative labels for some groups. Two or more entirely different approaches to the information architecture can be tested.

Deliverables

The deliverables from this technique are:

- a list of groupings and labels that are easily understood by users
- a list of groupings and labels that were difficult for users to understand
- ideas about location of content in the information architecture
- an understanding of tasks that users approach in different ways

Resources required

The evaluation is conducted by a facilitator and involves participants who are part of the user group for the intranet.

The staff resources needed include:

- a staff member to prepare materials and facilitate the evaluation
- a staff member to take notes during the evaluation
- end users to participate in the evaluation.

Number of participants

The number of participants required will depend upon the size of the information architecture and the range of tasks that staff need to undertake using the intranet.

When determining the number of participants, consider:

- Participants generally spend 10–15 minutes on this activity and may get through 10 scenarios in that time.
- Consistent responses appear after each scenario has been completed 4–5 times
- For an activity with 20 scenarios, 8–10 participants will be needed (each completing 10 tasks)

As with all usability testing, it is better to undertake a series of small tests (with changes in between tests) rather than one large test.

> **Important note**
>
> It is critical that the activity participants are actual end users of the intranet. There is little value in running the sessions with senior project stakeholders, as their way of thinking is unlikely to reflect the day-to-day users of the intranet.

Physical resources

The evaluation requires:

- index or filing cards sized about 76 x 127 mm
- pen or mailing labels to transfer the classification and scenarios onto the cards

Location

As the evaluation will only take 10–15 minutes of each participant's time, it is generally easiest to do it in their normal environment. If this is not practical, the evaluation can be held in a meeting room or office.

Analysis of results

Once the sessions have been completed, a staff member is required to document and analyse the results.

Time required

Tree testing is a rapid, flexible technique that can be used even when time is limited.

Table 1 shows the typical effort in person days, with high and low estimates for each step. Note that these estimates are provided as a guideline only and will depend upon a range of factors, such as:

- size of the information architecture
- range of scenarios to be included in the session

Step	Low estimate	High estimate
Prepare for evaluation • create list of tasks • create cards • select and schedule participants	0.5 day	1 day
Run evaluation	0.5 day	1 day
Analyse results	0.5 day	1 day
Write report	0.5 day	1 day
Totals	**2 days**	**4 days**

Table 1: Typical effort for tree testing (person days)

Undertaking tree testing

The major steps involved in undertaking a card-based classification include:

1. Create classification cards
2. Create scenario cards
3. Select and schedule participants
4. Run sessions
5. Analyse results

Each of these steps is described in the sections below.

1. Create classification cards

The first step is to create a set of index cards with the classification listed on them.

Before starting, number the classification using the following numbering system:

1. Top level item
    1.1. Second level item (section 1)
        1.1.1. Third level item (section 1.1)
    1.2. Second level item (section 1)
2. Top level item
    2.1. Second level item (section 2)
    2.2. Second level item (section 2)
3. Top level item
... and so on

Write the classification on index cards:

1. On the first card, write the items for the top level of the classification with their number. If there are more than can easily fit on a card, continue on another card.

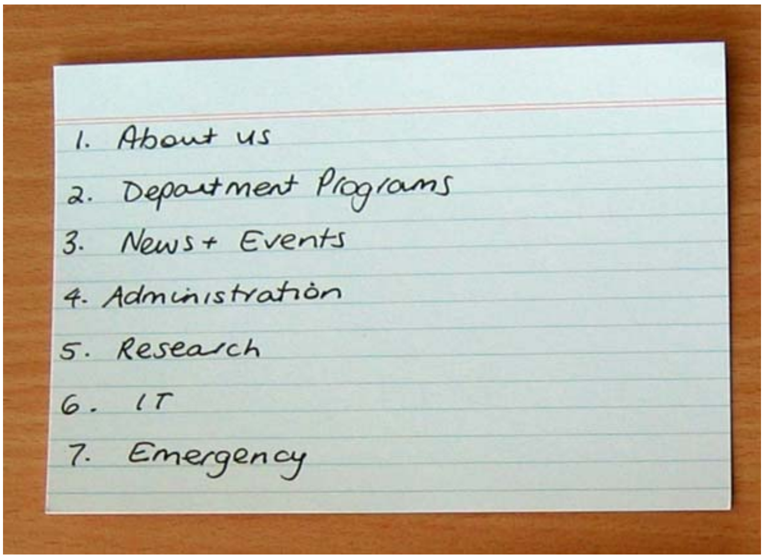

Figure 2: First level of the classification

2. On a new card, write the items for the second level of section 1 with their number (again, if there are more than will fit on a card, continue on another card)

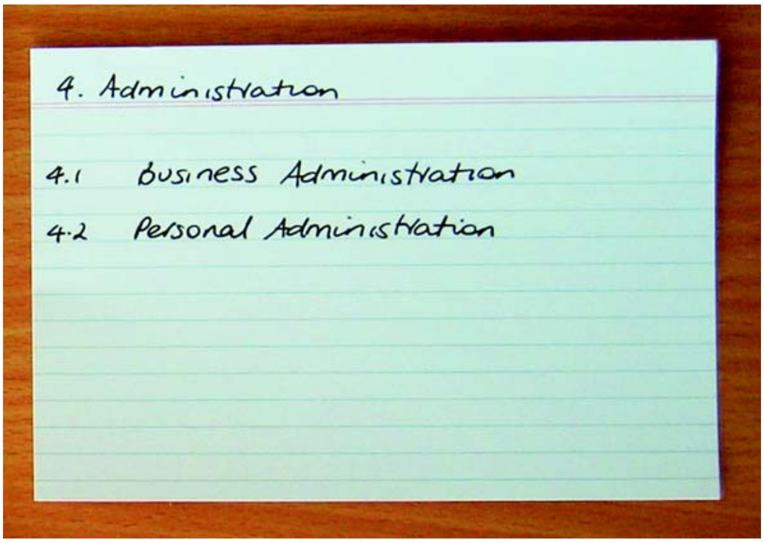

Figure 3: Second level of the classification

3. Continue until the entire classification is listed.

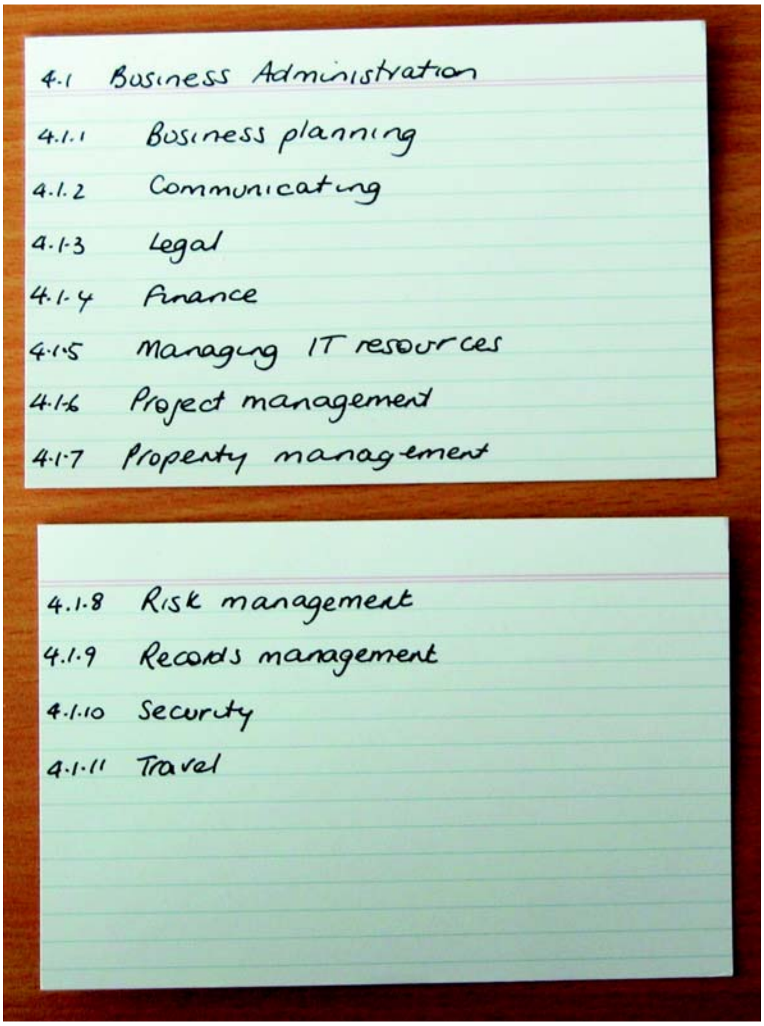

Figure 4: Third level of the classification

4. Bundle each section with an elastic band for ease of handling.

Although it may be time consuming, it is best to prepare the classification cards by hand. This allows more flexibility in the evaluation as it is possible to change the classification or labels during the evaluation if needed.

2. Create scenario cards

The scenarios to be used for the evaluation should:

- represent key tasks that people need to achieve using the intranet
- cover many areas of the classification

Write each scenario that will be used for the evaluation on an index card in clear and readable handwriting. Alternatively, print mailing labels that can be stuck onto the index cards. This is particularly useful if you need to create multiple sets of cards as it will save time writing.

Identify each card with a letter. This helps to identify individual scenarios when analysing the results. Keep the letters small so they do not distract participants from the actual scenario titles.

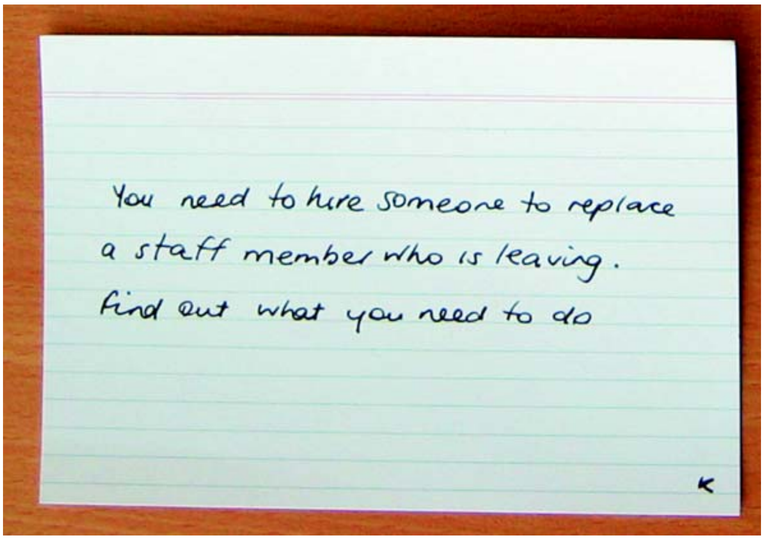

Figure 5: Scenario card

3. Select and schedule participants

Decide on the number of participants required for the evaluation. For guidelines, see *Resources required* (page 4). To ensure the results reflect actual user needs and ways of thinking, it is crucial that participants are, or will be, end users of the intranet.

When inviting participants, tell them they will be performing a simple exercise to provide some feedback on the progress of the intranet design or redesign. Let them know that it will only take 10–15 minutes and that there is no preparation required.

4. Run sessions

Introduce the session

To introduce the session, first introduce the facilitator and note-taker. Let the participant know that you:

- are working on a project to design or redesign the intranet
- have drafted some categories that might be used on the intranet
- would like to make sure that the categories are sensible for people who will be using the intranet.

Tell the participant that you will be asking them to tell you where they think they would look for information in the classification. Again, let them know also that it is a simple activity and will only take 10–15 minutes of their time.

A sample introduction may be:

First of all, we'd like to thank you for helping out. We've been working on a redesign of the Department intranet. We have talked with a lot of people about how the intranet can help them, looked at the content that we have, and come up with a draft classification that shows the main groupings for the site.

Before we start to build this on the computer, we would like to make sure that the groupings make sense and that people will be able to find information in the groups.

I'm going to show you a set of scenarios that people often use the intranet for [point to the scenario cards] and will ask you to show me where you think you would look in the classification [point to the hierarchy cards]. I'll then show you the next level of the classification.

If you look somewhere and it doesn't seem to make sense, let me know and you can look somewhere else. We'll only try 2 places – you don't have to hunt down the information. I want to make sure that the information is where you think it should be.

Work through the scenarios

Give the participant the pile of scenarios and let them manage those cards.

With the classification cards:

- Place the card(s) for the top level on the table.
- Ask the participant to read the scenario and tell you where in the classification they would first look for that information.
- For the item they have chosen, place the cards for the next level on the table (e.g. if the participant chooses item 1, show the card(s) that has 1.1, 1.2 etc listed).
- Continue until the lowest level of the classification is shown
- The note taker should write down the letter on the scenario card and the number for the lowest level item selected (e.g. scenario A, item 1.2.1.)
- Bundle the cards back up, but leave the cards for the top level on the table

Ask the participant to look at the next scenario and repeat the process until they have worked for around 10–15 minutes.

Figure 6: The hierarchy and scenario cards

If the participant looks at a card and does not think that they are in the right place, ask them if they would like to make another choice. Go back a level or go back to the beginning and follow their second choice. Do not ask them to look in more than two places as the intent is not for them to hunt through the classification – the intent is to gain an understanding of where they would look first.

For each scenario, the note taker should write down the letter on the scenario card and the number for the lowest level item selected (e.g. scenario A, item 1.2.1.). If the participant chooses twice, make a note of both paths. The note taker should also write down any interesting comments that the participant makes.

Handling the cards is the most difficult part of this process. Practice ahead of time so you can determine the best way to manage the cards without getting jumbled.

Wrap up the session

When the participant has worked for the allotted time, or is starting to look tired, wrap up the evaluation. Thank the participant for their help and ask them if they have any comments about the classification. The note taker should make a note of these comments.

Next participant

When starting with a new participant, start the scenarios from where you finished the previous session to cover all scenarios.

5. Analyse results

The aim of the analysis of results is to identify categories and labels that were suitable for the scenarios and those that were problematic. Analysis is simple and can be done with a spreadsheet.

In the first column of the spreadsheet, list out the classification (list the numbering system and the title). In the first row of the spreadsheet, list out the scenarios (list the letter and title for each scenario).

For each scenario undertaken, make a mark at the intersection of the scenario and classification item chosen. Use an upper case letter for first choices, a lower case letter for second choices. This will quickly show where people looked for each scenario (by looking down the columns) and what people expected to see in each part of the classification (by looking along the rows). Parts of the classification that worked well will show clusters of responses. For parts of the classification that didn't work well, the responses will be more scattered.

As an example, the spreadsheet in Figure 7 shows a good match between the scenario related to a mentoring program and the category of 'learning & development'. However, for the scenario related to travel allowance, participants suggested that they would look in a number of categories.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1			Hire new staff member	Carers leave	Salary package super	Rules for writing to clients	Travel allowanc e	Cab-charge	Get a new name badge	Title for a paper file	Mentoring program	Report a broken printer	Software purchase	Car parking
2			K	L	M	N	O	P	Q	R	S	T	U	V
3	Administration	4												
4	Business administration	4.1					X					X		
5	Business planning	4.1.1												
6	Communicating	4.1.2				XXXX								
7	Legal	4.1.3				XX								
8	Finance	4.1.4										X	X	
9	Managing IT resources	4.1.5										XX	XXX	
10	Project management	4.1.6			X									
11	Property management	4.1.7							X			X	X	XXXX
12	Risk management	4.1.8												
13	Records management	4.1.9								XXXX				
14	Security	4.1.10				x			X					X
15	Travel	4.1.11					Xx	XX						
16	Personal administration	4.2				X	x	X		X		X		
17	Certified agreement	4.2.1		XX	x									
18	Delegations	4.2.2					X							
19	Learning & development	4.2.3									XXXX			
20	OHS	4.2.4							X					
21	Orientation	4.2.5							X					
22	Managing staff	4.2.6												
23	Pay & conditions	4.2.7		XXXX	XXX		XX	X						
24	Performance management	4.2.8												
25	Personal safety	4.2.9						X						
26	Recruitment	4.2.10	XXX	X	X									

Figure 7: Example analysis spreadsheet

Look also at any comments that participants made during the session and consider what they mean for the classification.

There will be some parts of the classification or scenarios that were consistently chosen and some that were scattered. Consider the reasons for the results. A grouping may have worked well because:

- it was labelled well
- the scenario was straightforward
- the grouping is similar to an existing grouping

A group may have worked poorly because:

- the scenario may have been difficult to interpret
- the scenario may not be normally completed by the participant so they didn't know enough to know where to look
- the labelling may have been poor
- the grouping may have been poor

Guidelines and suggestions

Creating the scenario list

Generating the list of scenarios is a crucial part of preparing for the evaluation. When selecting scenarios, consider the following:

- Make sure the scenarios cover a wide range of the classification in order to ensure many areas of the classification are tested.
- Ensure that there are not too many scenarios in one part of the classification. Participants will naturally go back to parts of the classification that they have already seen.
- The scenarios should be in the format of short stories, and should provide background where possible.
- Make sure the scenarios cover administrative and business related tasks.

Example scenarios

Examples of administrative scenarios that can be used for an intranet include:

- You've moved to a new desk. How do you get an ergonomic assessment of your desk
- How do you change a deduction from your salary.
- You have cut your hand – find out where your nearest first aid officer is
- You need to get a statutory declaration signed. Find out if there are any Justices of the Peace nearby
- You were in a meeting where someone made an inappropriate comment – find out who to talk to about it
- Your boss has asked you to develop a risk assessment for your project. Find out more.
- Your password has expired. Find out how to get it changed.
- Find out whether you can salary package extra money into superannuation

Involving staff with disabilities

This technique can involve staff with disabilities and is a valuable way of ensuring that a wide range of staff are included in project activities:

- Staff with occupational overuse injuries can be involved in the same way as other participants.
- For staff with hearing impairments, print out the instructions and follow the normal process for the rest of the activity.
- For staff with vision impairments, read out the scenarios and the choices. This is a valuable way to determine whether the list of choices is too long for staff who are using screen reader software.

Variations on the technique

Evaluating non-hierarchical classifications

Although the technique as described here is applied to hierarchical classifications, it can also be applied to a non-hierarchical classification. In this situation, it can be used to evaluate a list of options such as the main items that will appear on the home page.

Assessing multiple hierarchies

More than one hierarchy can be used during the evaluation. For example, an intranet may provide both a subject and a task-based classification scheme. This can be explored in a number of different ways:

- By showing participants both classification schemes at the same time, this technique can be used to identify the types of tasks people would use with each classification scheme.
- By evaluating the classification schemes separately, the technique can show which is most likely to be successful at allowing staff to find information.

Evaluating other classification schemes

This technique can be applied to classifications other than those used as an intranet information architecture. For example, card-based classification could be applied to:

- records management classification schemes
- taxonomies
- keyword lists

In all these cases, the basic approach remains the same, with the tasks simply adapted to fit the classification scheme being evaluated. For example, in a records management classification, participants could be shown a document and asked how they would classify it; while for a keyword list, participants would be asked which keywords they would use to describe a given document.

Using a page layout instead of index cards

This technique can be used to test a page layout. Rather than writing the classification on index cards, print out the existing or planned pages and use them. The disadvantage of this approach is that for a large site, many pages may be needed (it is easier to handle a set of index cards than a printed set of screens).

Using as part of a focus group or staff interview

This technique can be applied as part of a focus group or staff interview in the early stages of a design or redesign project. This can be a useful way to make these activities more practical.

Applying the results

The outcome of this technique is primarily used to improve a classification such as the intranet site structure. The results help to:

- identify parts of the classification that are good and those that need improvement
- provide additional information on labelling ideas
- suggest placement for content in the classification

The results from the activity should be incorporated into the classification and a new evaluation should be run (with new or the same scenarios) until the classification is considered to be usable enough to proceed to the next stage.

Case study

This case study is based on tree testing conducted during the redesign of a Government Agency intranet.

Research conducted prior to the evaluation

The evaluation was conducted as soon as the intranet's high-level information architecture had been drafted. The information architecture had been drafted based on user research activities such as focus groups, stakeholder interviews and card sorting.

The user research provided a good understanding of which tasks the intranet could help people with. It also provided a large set of scenarios for the evaluation.

Recruiting participants

In earlier stages of this project, it had been hard to get some areas of the Agency involved in the research activities; most people weren't very interested in the intranet and it was very hard to schedule some teams.

Rather than arrange participants ahead of time, it was arranged with some managers to be at their location on a particular afternoon. During that time, staff were approached at their desks and asked to provide 10 minutes of their time to do an exercise. Most people thought that this was a bit strange, but 10 minutes didn't seem like a lot, so they were happy to be involved.

This flexible approach included staff who were otherwise difficult to involve in the project, and made sure we included a wide range of people in the testing.

Running the session

Altogether, there were 24 participants across four locations (six in each location over a two hour period). The evaluation was conducted initially in two locations, and the results analysed in a spreadsheet. Minor changes were made to the classification, including changing several labels. The evaluation was then conducted in two other locations.

The user research and card sorting sessions had provided good input into the development of the classification, meaning that only minor changes were needed to ensure that staff could easily find information to complete their tasks.

Further evaluation

Once the site was built, further usability testing was conducted on the page layouts and content, further testing the information architecture. A different set of scenarios were used, and it was found that staff were able to easily find information to complete the scenarios.

Authors of this Guide

While the Project Guides are reviewed by the whole Step Two Designs consulting team, the following team members (present or past) were the primary contributors to this Guide:

- Donna Maurer