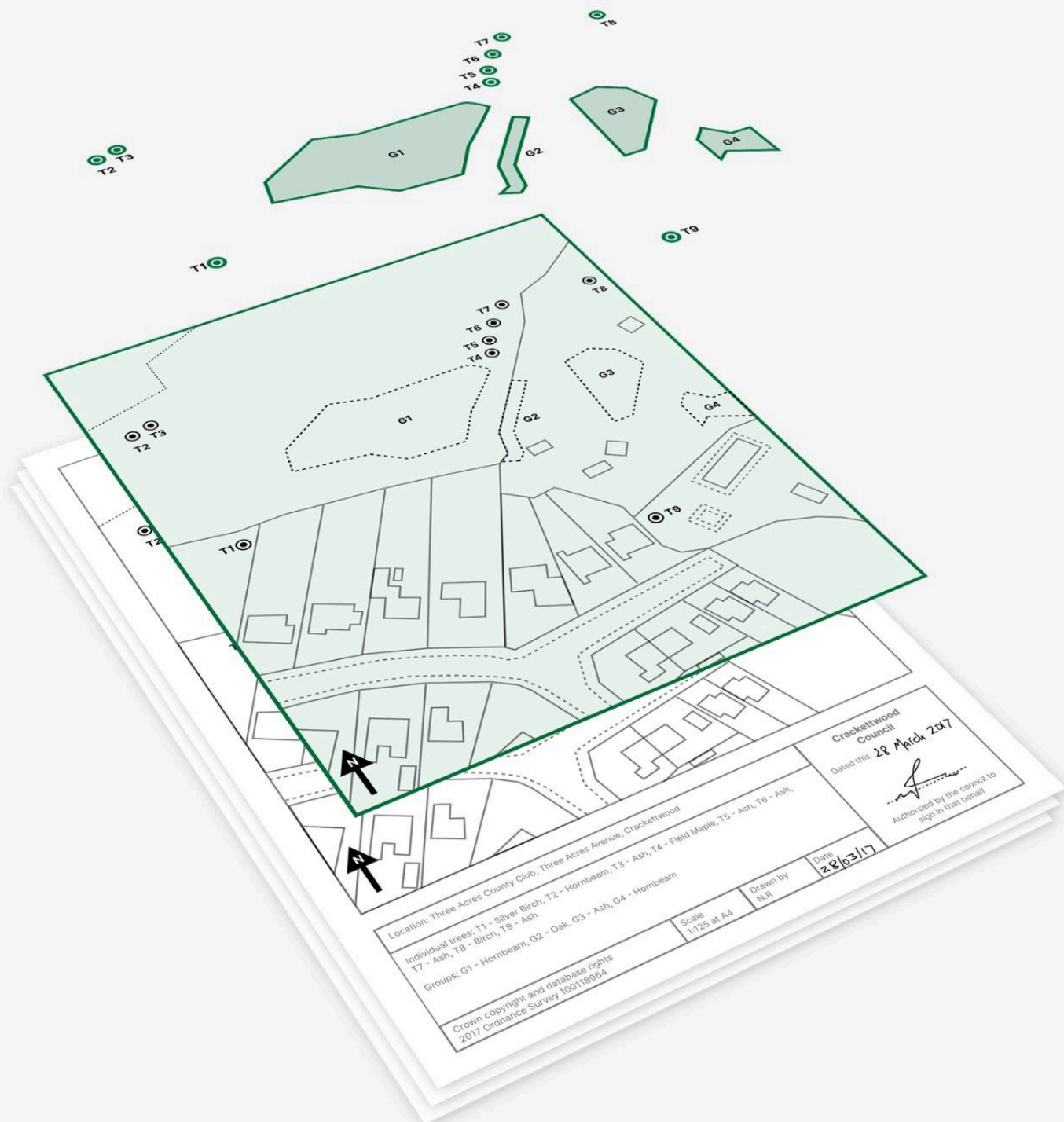


EXTRACT

Powering digital planning by
converting planning documents
into geospatial data



Extract Evaluation

Abstract

To evaluate Extract, we conducted user research with over 30 local planning authorities and analysed over 400 digital planning records.

We found that Extract processes documents quickly and often produces a good draft of a digital record, however it is not suitable to be an automated solution. User review and correction is essential to produce good quality records.

We estimate that Extract can reduce the total time local authorities spend creating digital records by around a half.

Executive summary

Extract uses artificial intelligence (AI) to create digital planning records for the following three planning document types:

- Conservation Areas
- Article 4 Directions
- Tree Preservation Orders

Users upload scans or photographs of the original document, and can review and edit the AI outputs before exporting their finished records.

To evaluate Extract, we conducted user research with over 30 local planning authorities and analysed over 400 digital planning records.

This report describes our conclusions on Extract's accuracy and potential benefits.

Key findings

1. We estimate that there are 190,000 - 240,000 Conservation Area, Article 4 Direction and Tree Preservation Order documents in England. Of these, so far only 20 - 25% have been published on MHCLG's Planning Data Platform. (See [Planning records](#))
2. Planning documents can be complex and creating digital records is a time-consuming task. We estimate that, with existing tools, creating a

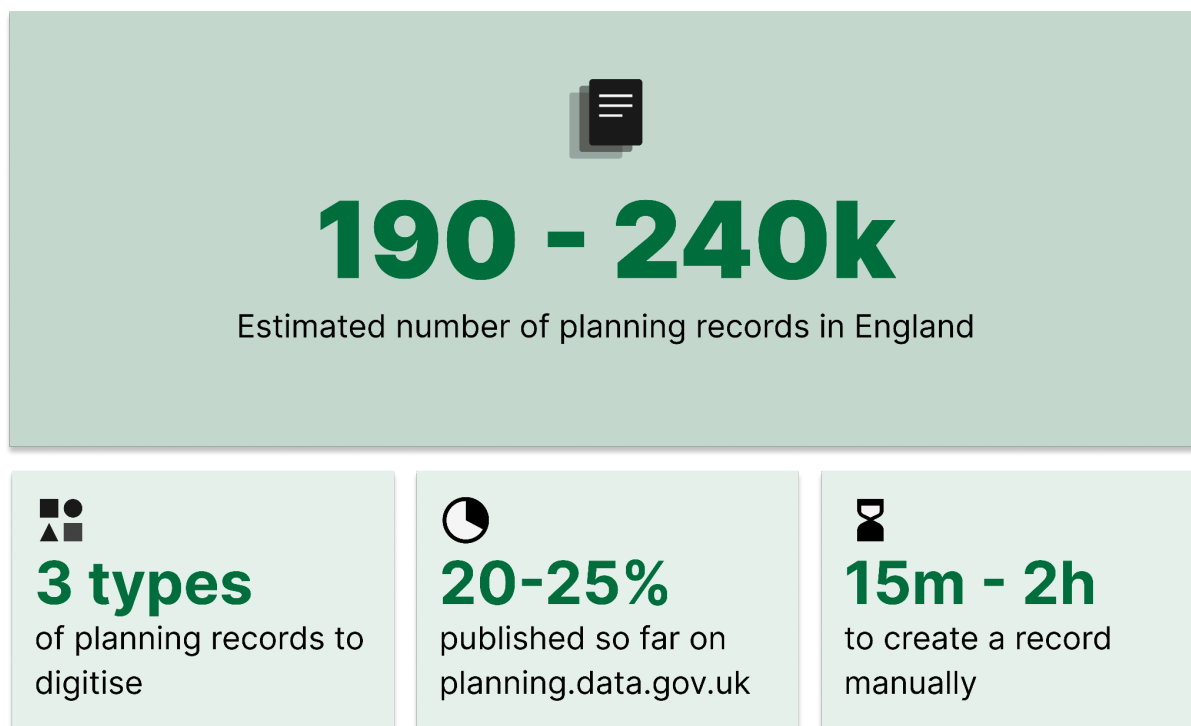
digital record takes between 15 mins and 2 hours. (See [How local authorities currently create digital records](#))

3. Extract processes documents quickly. On average it takes two minutes to produce a draft digital record. (See [Processing time](#))
4. Extract often produces a good draft of a digital record, however it is not suitable to be an automated solution. User review and correction is essential to produce good quality records. We expect that Extract's outputs will require only minor edits in around two-thirds (59 - 72%) of cases. (See [Geospatial accuracy](#))
5. Extract's accuracy rate varies by document type, primarily due to the age of documents and the locations they are typically used. (See [Geospatial accuracy](#))
6. Since AI is non-deterministic, Extract can produce different outputs for the same input. When we ran Extract over our test set three times, we found that it produced at least one good quality output for three-quarters of documents. (See [Non-determinism and consistency](#))
7. When Extract's geospatial outputs are accurate, it significantly reduces the time taken to produce a digital record. We estimate it offers time savings of 60 – 80%. (See [Time savings](#))
8. For the remaining cases, Extract's outputs are only partially accurate. We expect that some elements will still be useful, but the user will need more time to produce a good quality record. (See [Time savings](#))
9. Overall, we estimate that Extract can reduce the total time local authorities spend creating digital records by around a half (40 - 55%). (See [Time savings](#))

Introduction

The scale of the digitisation challenge

There are an estimated 190,000 – 240,000 Conservation Area, Article 4 Direction and Tree Preservation Order documents in England. Around one-fifth have been published to the Planning Data Platform so far.



Background

The Incubator for Artificial Intelligence (i.AI)

We are the government's expert applied AI team and we design, test, and deliver AI products for the public sector.

The Digital Planning programme

The Digital Planning programme within Ministry of Housing, Communities & Local Government (MHCLG) has a mission to modernise the planning system and make it fit for the 21st century. We enable more informed and more effective plan and decision-making in England using a modern, data-driven approach to shaping communities and the built and natural environment.

At the heart of this transformation is a Planning Data Platform (planning.data.gov.uk) that makes key data easier to find, use and trust. By standardising planning data across England, and making the evidence accessible, we are laying the groundwork for a more efficient system. This data allows local planning authorities to adopt user-friendly planning software that is already showing impressive potential.

Extract

To accelerate local authorities' digital planning transformation, MHCLG has partnered with i.AI to create Extract: an AI tool to help local planning authorities turn planning documents into data faster, more consistently, and at lower cost than manual methods.

Extract uses AI to automatically create digital records from scanned or photographed originals, which users can then review and edit before publication.

Our evaluation

The aim of this evaluation is to understand the performance and potential benefits of Extract. In particular, we aim to answer the following questions:

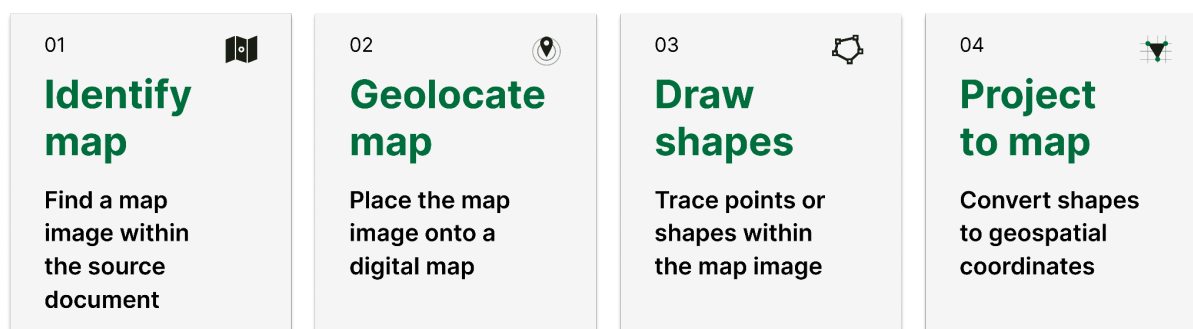
- How many documents need digitising across the planning system?
- How accurate are Extract's outputs? How does Extract's accuracy compare to human performance?
- How long does it take a user to digitise a document using Extract? How much faster is it than using current methods?

The evidence base for this evaluation is:

- User research, conducted with over 30 local planning authorities over 6 months
- Analysis of Extract's accuracy over a representative test set of over 400 planning documents, where its output can be compared to manually generated outputs.

What does Extract do?

Extract creates digital records by capturing text, dates and geospatial data from planning documents. The process by which it captures geospatial data mirrors the process by which a GIS officer would complete the same task.



Planning records

Extract can create digital records for three types of planning documents:

- Conservation Areas
- Article 4 Directions
- Tree Preservation Orders

Based on what is available on MHCLG's Planning Data Platform so far, we estimate that there are around 12,000 Conservation Areas in England and 9,000 – 16,000 Article 4 Directions. There are a larger number of Tree Preservation Orders: we estimate around 200,000.

Table 1: Estimated number of planning records in England by type, and the proportion already published to MHCLG's Planning Data Platform (90% confidence intervals)

	Conservation Areas	Article 4 Directions	Tree Preservation Orders	All document types
Average number per local authority	32 - 39	25 - 50	490 - 630	560 - 680
Total number in England	12,000 - 12,400	9,000 - 16,000	175,000 - 215,000	190,000 - 240,000
Proportion on the Planning Data Platform	86 - 88%	18 - 30%	16 - 20%	20 - 25%

Key finding 1: We estimate that there are 190,000 - 240,000 Conservation Area, Article 4 Direction and Tree Preservation Order documents in England. Of these, so far only 20 - 25% have been published on MHCLG's Planning Data Platform.

Local authorities can also make their data available online in other formats. From reviewing a small sample, we estimate that the proportion of local authorities that have digital records available online is as follows.

Table 2: Estimated proportion of local authorities with digital records available as open data by document type

Document type	Local authorities with open data
Conservation Areas	90 - 100%
Article 4 Directions	50 - 85%
Tree Preservation Orders	40 - 75%

How local authorities currently create digital records

Creating a digital planning record includes capturing:

- details of the record itself, for example the title of the order, the date when it came into force and the date it was published
- the geographical area or location to which the record applies, i.e., plotting affected properties and trees on a map
- details of the areas and locations affected, e.g., area names, addresses, tree species

The complexity of planning documents can vary widely: from an Article 4 Direction affecting a single property, to a Tree Preservation Order protecting several hundred individual trees.

Different local authorities store their planning records in different formats and systems. However, to take advantage of the tools and services developed by MHCLG's Digital Planning programme, they must also publish it in a [standardised schema](#).

During our user research, we observed users creating digital records with existing processes and tools. The process is laborious and prone to error. Furthermore, when the geographical area affected by the order is complex, specialist GIS tools and skills are required.

Key finding 2: Planning documents can be complex and creating digital records is a time-consuming task. We estimate that, with existing tools, creating a digital record takes between 15 mins and 2 hours.

Creating digital records with Extract

Extract is a 'human in the loop' AI system, which means that user review and correction is part of the workflow. The process of creating a digital record with Extract is as follows:

1. Users start with a digital scan or image of the original document.
2. Users select the document type they would like to digitise, and upload the scan. The document is processed using AI to produce a draft digital record.
3. Users review the output and check it against the original. Extract includes a variety of editing tools, and a pop-up window to view the

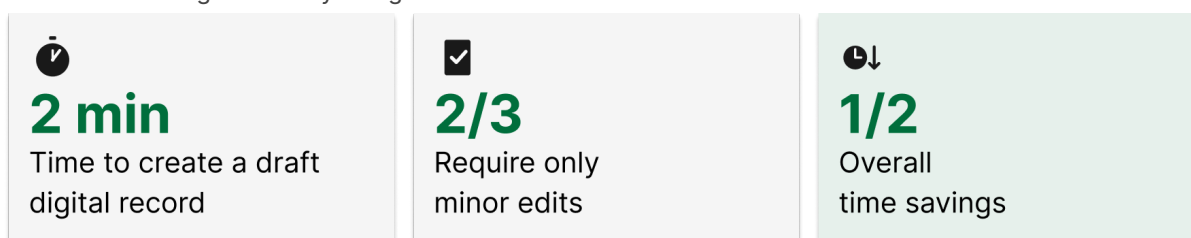
original document. A workflow invites users to review and correct each element of the extraction in turn, before marking it as complete.

4. Users download their finished records. Records are available in common geospatial file formats, ready to import into local authority systems. Users can also download records in bulk, for publishing to the Planning Data Platform.

Extract's performance

Fast, often good - but human review is essential

Extract produces draft records in minutes and most are usable with only minor edits. We estimate that the overall time saving offered by using Extract over current manual methods is around one-half.



To understand Extract's accuracy, we tested it on over 400 existing digital records, where Extract's output can be compared directly to manually generated outputs.

For each document type, we selected a representative sample of 10 local authorities based on their region and rural urban classification. Then we collected digital records for those local authorities.

The composition of the resulting test set was as follows.

Table 3: Number of documents in our test set by document type

Document type	Number of documents
Conservation Area	211
Article 4 Direction	80
Tree Preservation Order	122
All document types	413

The documents we collected for Conservation Areas were typically Conservation Area appraisals, whereas for Article 4 Directions and Tree Preservation Orders we collected scans of the original documents. This means

that few Conservation Area documents in our test set are “historic”, relative to Article 4 Directions and Tree Preservation Orders.

We analysed Extract’s performance over this representative test set, then used statistical techniques to estimate Extract’s expected performance over *all* planning documents.

Since Extract’s output includes a variety of data types – including text, dates, categorical data and geospatial data – we measure the accuracy of each type of output and each document type separately. In doing so, we give more weight to geospatial outputs, as they are more challenging to generate, review and correct.

Because AI is non-deterministic, we also tested the consistency of Extract’s outputs over several runs.

Processing time

We tracked the run time of Extractions during our testing. Most extractions took between one and three minutes, but some took as long as ten minutes.

Key finding 3: Extract processes documents quickly. On average it takes two minutes to produce a draft digital record.

Geospatial accuracy

We define the accuracy of geospatial outputs by document type.

A good quality extraction of a Conservation Area or Article 4 Direction area has at least 80% overlap with the original, manually-drawn area. This criteria ensures that both the shape and the position of the area is a good match.

Figure 1 shows an example of an Article 4 Direction area extraction which was just above the threshold for ‘good quality’.

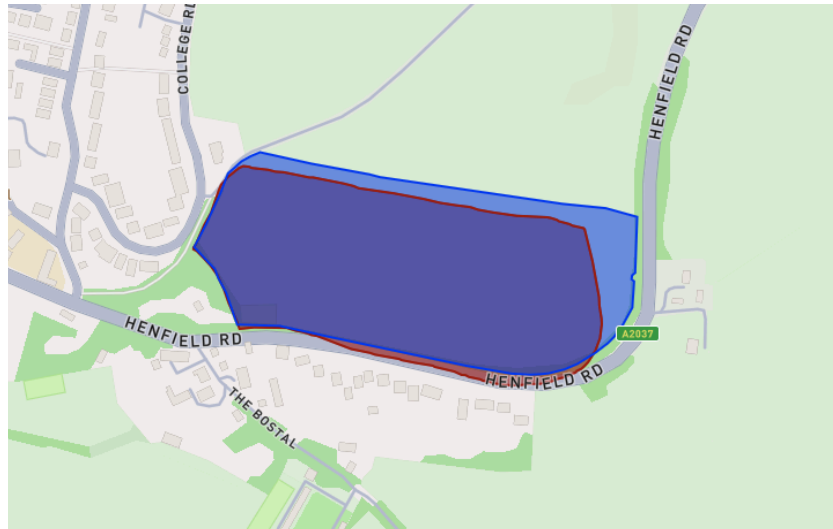


Figure 1: Article 4 relating to land at Mains Farm, Upper Beeding. The area from the original digital record is shown in blue, and Extract's attempt is shown in red. The shape of the Article 4 Area in Extract's output is correct but it has been slightly squashed and misplaced. This is an example of an area extraction which was just above the threshold for 'good quality'.

A good quality extraction of a Tree Preservation Order has at least 80% of zones (groups, areas or woodlands) and trees close to the right place. This criteria ensures that the map has been geolocated well, although some trees and zones may be misshapen or slightly misplaced. Figure 2 shows an example of a Tree Preservation Order extraction which was just above the threshold for 'good quality'.

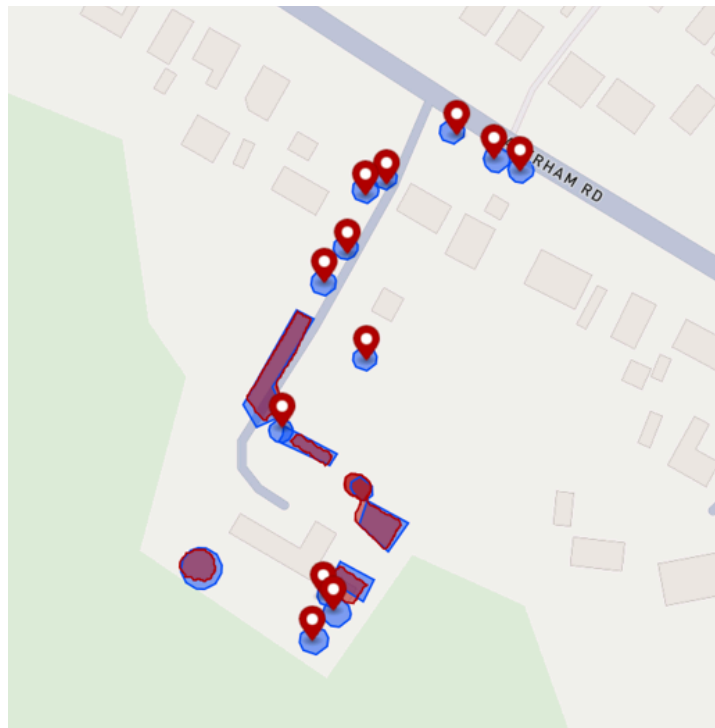


Figure 2: The Broadland District Council Tree Preservation Order Taverham 2006 No.9. The original digital record is shown in blue, and Extract's attempt is shown in red. Note that, in the centre of the image, Extract has merged two adjacent groups that were disjoint in the original record. This is an example of a Tree Preservation Order whose geospatial extraction was just above the threshold for 'good quality'.

Our threshold for 'good quality' doesn't imply that Extract's outputs are completely correct. Instead, we aim to capture those instances where users only have to make minor edits.

We find that Extract's geospatial outputs are often good quality, but there are also instances where one or more steps in the extraction process performs poorly.

Key finding 4: Extract often produces a good first draft of a digital record, however it is not suitable to be an automated solution. User review and correction is essential to produce good quality records. We expect that Extract's outputs will require only minor edits for around two thirds (59% to 72%) of cases.

We discuss common issues that we have observed in detail below (see [Common errors](#)), but the main factors are:

- The document age – since an area may have changed significantly over time, and the style of older maps may be hard to match to modern maps,
- Whether the area is rural or urban – since there are more landmarks in urban areas, which help Extract to identify the correct location.

Key finding 5: Extract's accuracy rate varies by document type, primarily due to the age of documents and the locations they are typically used.

Extract's performance on each document type is as follows.

Conservation Areas

For Conservation Areas, we assessed Extract's performance against the following criteria.

Table 4: Our assessment criteria for Conservation Areas and Extract's expected success rates

Criteria	Extract's estimated success rate
The Conservation Area should overlap with the original area by at least 80%	81 - 89%

We expect the Conservation Area to have at least an 80% overlap (intersection over union) with the manually-drawn area. We aim for a higher overlap threshold for Article 4 Direction areas and Conservation Areas as they are often larger and more complex areas, making them more time consuming to edit at the user review stage.

Extract performs well on Conservation Areas as Conservation Area appraisals are often newer documents, so the maps in them are easier to match to a modern map.

Article 4 Directions

For Article 4 Directions, we assessed Extract's performance against the following criteria.

Table 5: Our assessment criteria for Article 4 Directions and Extract's expected success rates

Criteria	Extract's estimated success rate
Permitted development rights should be at least 80% correct	69 - 88%
The Article 4 Direction should overlap with the original record by at least 80%	40 - 58%

Extracting the permitted development rights which are affected by an Article 4 Direction is a multiple-choice selection. So, by 80% correct, we mean that at least 80% of Extract's choices match the choices in the original record (precision) and vice versa (recall). This ensures that Extract neither misses out nor falsely includes too many development rights.

Our confidence in this assessment is not as high as for other criteria, since permitted development rights were often not available in the digital records we sourced, so we had fewer examples to test with.

The second criterion concerns the mapping of the Article 4 direction area, and is identical to the criteria for Conservation Areas.

Extract doesn't perform as well on Article 4 Direction areas as it does on Conservation Areas since:

- the original documents are often older
- Article 4 Directions can apply to land, as well as buildings, and maps of these areas often have fewer of the landmarks that Extract uses to geolocate the map.

Tree Preservation Orders (TPOs)

For Tree Preservation Orders, we assessed Extract's performance against the following criteria.

Table 6: Our assessment criteria for Tree Preservation Orders and Extract's expected success rates

Criteria	Extract's estimated success rate
All trees and zones from the original record should be included	90 - 97%
At least 80% of the trees and zones should be mapped close to the right place	58 - 72%

The first criterion concerns the metadata of the individual entities (trees, groups, areas and woodlands) in the TPO, e.g., "T1 - Sycamore". We expect all entities to be present in the TPO record, regardless of whether they are mapped correctly. This criterion ensures that no entities are missed out in the user review stage, and details such as tree species - which are easy for Extract to capture but slow for users to capture manually - are included.

We found a few examples where Extract included extra entities, which weren't in the original record - these were not hallucinations, but examples of when a tree had been felled or a zone had been removed from the order.

This second criterion concerns the mapping of the TPO. We expect at least 80% of the original entities to be mapped in roughly the right place: for trees this means within 10m of the correct location, and for zones this means at least a 30% overlap (intersection over union) with the correct area.

We aim for a smaller overlap threshold for Tree Preservation Order zones as they are typically small areas and simple shapes, making them both harder for Extract to position accurately and easier for users to edit and correct.

Common errors

Extract typically captures geospatial data from planning documents in three steps. These steps mirror the manual process by which a GIS officer would complete the same task. Each step tends to fail for different reasons.

As steps 2 and 3 are performed independently, if Extract performs poorly in one step, the output of the other is often still useful. Extract gives users the tools to verify and correct the outputs of each step in turn.

Step 1: Locating a map within the document

The first step is to locate a map in the planning document and extract the map as an image.



Figure 3a: An example of a map image extracted from a planning document from Worthing Borough Council

If the user uploads a document of the correct type and that document contains a map then this step is almost always successful. Issues can occur when the source document does not contain a map or when a map has been split across multiple pages.

Step 2: Geolocating the map image on a digital map

The next step is to locate the map image and align it with a modern, digital map.

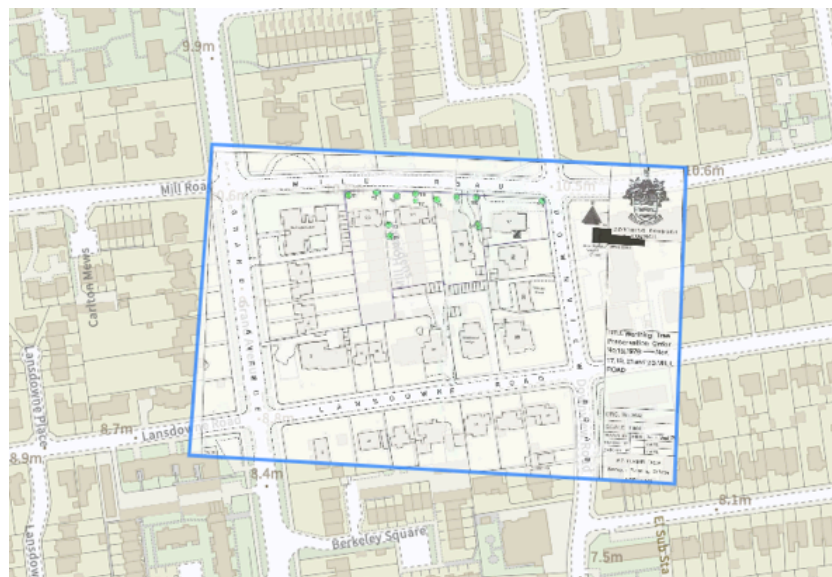


Figure 3b: The map image after geolocation

Extract relies on named landmarks within the map image to locate it. Hence it can struggle when these are not available, for example if a map is very large scale, is styled without road or building names, or is very faded or blurred.

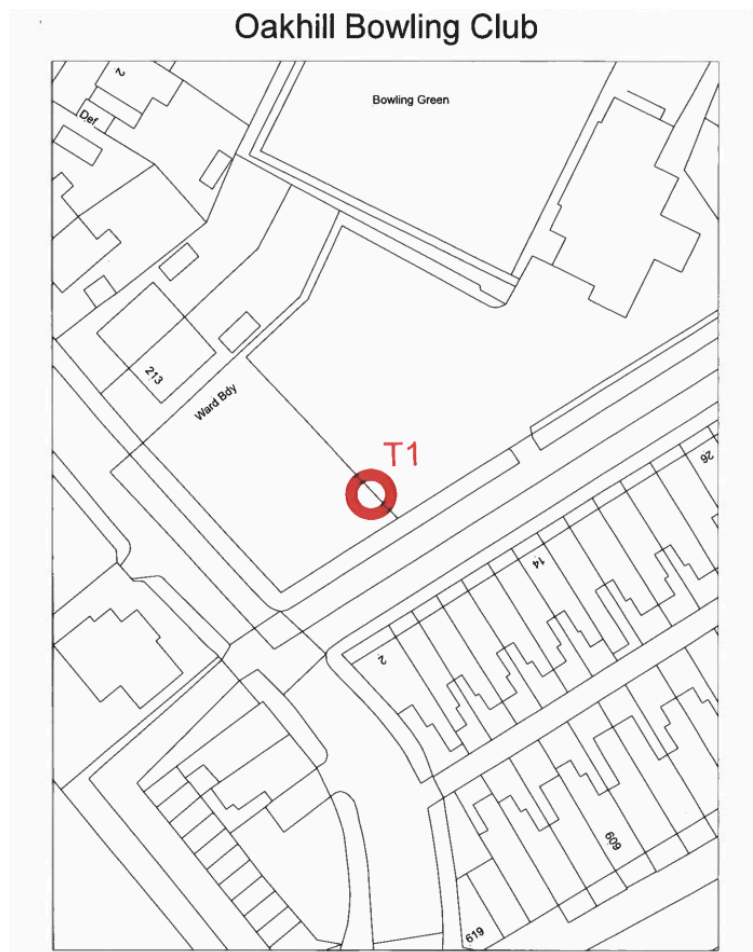


Figure 4: A map image where the geolocation step of Extract failed during our user testing as there are no road names, addresses or even local authority names printed on the map. Source: https://www.stoke.gov.uk/directory_record/335510/tpo_148_%E2%80%93_land_to_the_south-west

This step can also fail if Extract cannot match the map image to a current digital map. Once Extract has identified the locality of a map image, it uses the shape of map features such as buildings and roads to align the map image to the digital map. Extract can struggle if an area has changed significantly over time, if a map is very small scale (i.e., a single property), or if a map displays a rural area which doesn't contain many landmarks.

Step 3: Drawing the points or shapes of interest within the map image

The last step is to draw the shapes or points of interest within the map image, and project these onto the digital map to create geospatial areas and points.

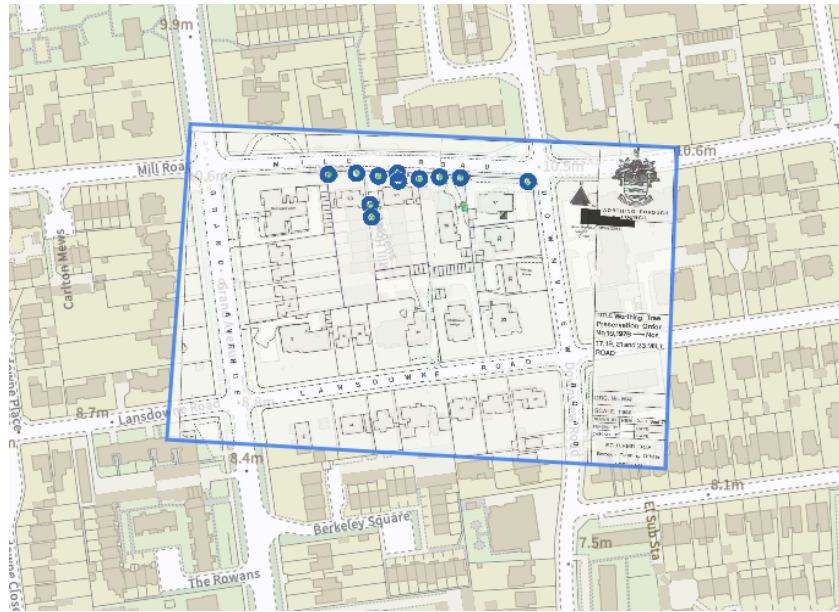


Figure 3c: The geolocated map image with the trees drawn on the map marked as points

In this step, Extract can struggle to properly trace areas on the map which are complex, crowded by other map features, or displayed with dotted lines or hatched areas. Area Townscape Appraisals are an example of maps which can cause issues at this stage, as they often plot many different features with varying colours and styles.

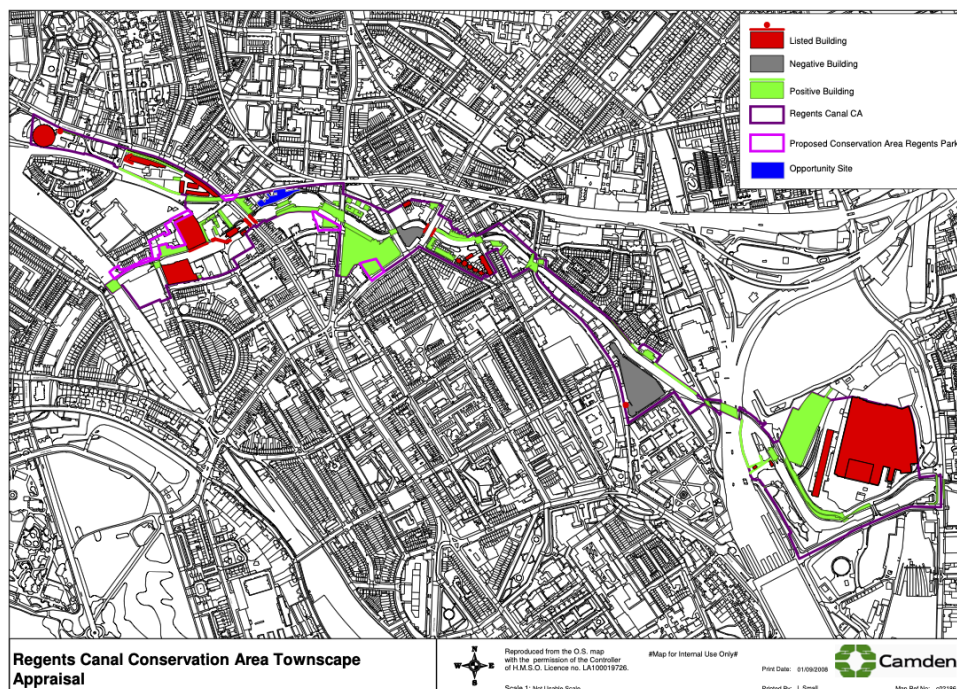


Figure 5: Regents Canal Conservation Area Townscape Appraisal, which plots two different Conservation Areas, alongside other map features which overlap with the Conservation Area boundary. In our testing, Extract failed to identify and accurately draw the Conservation Area boundary. Source: <https://www.camden.gov.uk/regent-s-canal-conservation-area-appraisal-and-management-strategy>

Extraction of free text and dates

Extract is very good at transcribing text and dates from source documents, including from hand written text. However we struggled to evaluate its accuracy over these fields, as in our test set they were often empty, and where they were completed they were not used consistently.

Document name

Among records on the planning data portal this field was often left empty (16% of records). Extract names records according to the title of the original document and these can be quite verbose. In user testing, users sometimes expressed a preference for a more succinct name or a reference number, e.g., "Tree Preservation Order E/27/99", rather than "Parish of Burrough Green in the County of Cambridgeshire Tree Preservation Order E/27/99 Land at The Old Post Office".

In this respect Extracts creates records which are more faithful to the original document, but that may not cater to local preferences.

Dates

Different document types have different dates which can be recorded. In manually generated records, date fields are often left empty. For example, 26% of records on MHCLG's Planning Data Platform had no start date.

Extract tends to complete more of these fields, but can only do so if dates are available in the source documents it is given. For example, the designation date of a Conservation Area was often not available in the Conservation Area appraisals we tested (or were quoted as, say, a year rather than a date).

We observed that end dates were almost universally empty, and the meaning of the start date field is ambiguous. Hence we test Extract's ability to capture dates with the made and confirmed dates of Tree Preservation Orders, which are unambiguous and often available in source documents.

We find that, where the source document allows, Extract captures dates more often and more accurately than for manually created records. Where dates are not available in the source document, Extract typically leaves the field blank but occasionally selects another (incorrect) date from the document. We did not observe any hallucinations, i.e., Extract citing dates that were not from the source document.

Table 7: The accuracy of Extract's date capture versus manually created records. Note that, if the date was not available in the source document, the accuracy of manually-created records could not be judged.

Source document	Date field in digital record	Made date		Confirmed date	
		Manually created	Extract	Manually created	Extract
Date was available	Correct	66%	93%	59%	74%
	Incorrect	1%	0%	7%	0%
	Could not be judged	3%		0%	
	Empty	27%	0%	9%	0%
Date was not available	Incorrect	-	2%	-	1%
	Empty	-	1%	-	25%

Non-determinism and consistency

AI is non-deterministic, meaning that the same input can produce different outputs on different occasions.

When we ran Extract over our test set three times, we found that for around one in five planning documents, Extract produced a good draft of a digital record on some occasions and a poor draft on others. While Extract produced consistently good quality outputs for two-thirds (64%) of documents, it produced at least one good output for three-quarters (78%) of documents.

Table 8: The consistency of Extract's outputs over three runs, by document type

Planning documents for which Extract's draft was...	Conservation Area	Article 4 Direction	Tree Preservation Order	All document types
...good on all three occasions	79%	39%	56%	64%
...sometimes good and sometimes poor	10%	19%	18%	14%
...poor on all three occasions	12%	42%	26%	22%

Key finding 6: Since AI is non-deterministic, Extract can produce different outputs for the same input. When we ran Extract over our test set three times,

we found that it produced at least one good quality output for three-quarters of documents.

Time savings

During our user research, we observed users creating digital records both with existing processes and tools, and with Extract. Based on this, we have estimated the time savings that Extract offers local planning authorities.

When Extract's geospatial outputs were good quality and only minor edits were needed, users completed their review and corrections in 5 – 20 minutes, depending on the complexity of the document and the nature of the corrections required.

Key finding 7: When Extract's geospatial outputs are accurate, it significantly reduces the time taken to produce a digital record. We estimate it offers time savings of 60 – 80%.

When Extract's geospatial output is not good quality, users can use its editing tools to make corrections. The majority of outputs will be partially accurate, but we do not have data on how long it takes users to make more complex corrections. For the purposes of this evaluation, we assume that the time needed to make complex corrections is the same as the time taken to digitise a document from scratch, but this is likely an over-estimate.

Key finding 8: For the remaining cases, Extract's outputs are only partially accurate. We expect that some elements will still be useful, but the user will need more time to produce a good quality record.

Key finding 9: Overall, we estimate that Extract can reduce the total time local authorities spend creating digital records by around a half (40 - 55%).

This means that, for a typical local authority with around 600 planning documents, we expect the time taken to create all of their digital records with Extract is 260 hours, versus 515 hours with manual methods.

Since these estimates are based on a small number of users, and in user-testing conditions rather than real-world usage, our confidence in these estimates is low.

Annex: Our evaluation methodology

Document number estimates

We estimated the number of planning documents in England by extrapolating from the records on MHCLG's Planning Data Platform.

We calculated the number of records per local authority and per document type, and used bootstrapping methods to estimate the distribution of the average number of records. We assumed that each local authority's dataset was complete, and used data cleaning methods to de-duplicate Tree Preservation Orders for some local authorities.

We used the average number of records per local authority and per document type to estimate the number of planning records that are yet to be published, assuming that those local authorities which have already published their data are representative of those who have not.

The total number of local authorities we expect to publish planning data is 348, which is the number of London Boroughs (incl. the City of London), Metropolitan Districts, Non-Metropolitan Districts, Unitary Authorities and National Park Authorities.

Extract's performance

To understand Extract's accuracy, we tested it on over 400 existing digital records, where Extract's output can be compared directly to manually generated outputs.

For each document type, we selected a representative sample of 10 local planning authorities based on their region and rural urban classification. Then we collected digital records for those local authorities. The method we used varied by document type:

- For Conservation Areas and Article 4 Directions, where digital records are more often available online, we sampled from all local authorities in England. We then collected all planning documents we could find. This means that local authorities in our sample that have not published data online, and who did not respond to direct requests, are under-represented in our test set.
- For Tree Preservation Orders, where existing digital records are more difficult to obtain, we sampled from local authorities which had published data on MHCLG's Planning Data Platform. Then, as Tree Preservation Orders are more numerous, we sampled 3-4% of each local authority's documents, stratifying the sample on the order's start date.

Note that our test set did not include planning documents from a National Park Authority, unless they were cross-posted on a local authority's website.

Note also that our test set comprised documents that are already available as digital records (either on MHCLG's Planning Data Platform or on the local authority's website), and included planning documents of any age (not just

"historic" documents). We assume that these documents are representative of the document set as a whole – in particular, that they are not significantly different from documents which have *not yet been digitised*, which is Extract's primary use case.

Time savings

We expect that the time taken to create a digital record - both using manual methods and using Extract - depends mostly on the complexity of the original document. Hence to estimate the time savings offered by Extract, we modelled the distributions of:

- The time taken to create a digital record manually, and
- The time saving factor of Extract.

We suppose that the time taken to create a digital record manually ranges between 15 minutes and two hours, with most taking around 40 minutes. Similarly we suppose that when Extract's outputs are good quality it offers a time saving factor ranging from 2 to 7 (equivalent to 50% to 85% time saving), with a factor of 4 (75% time saving) the most likely scenario.

Since these estimates are based on a small number of users, and in user-testing conditions rather than real-world usage, our confidence in these estimates is low.

Then, to model the total time taken to digitise a number of documents, we sample from these distributions and the distributions for the Extract performance rates of each document type.

When Extract's draft is not good, we assume that the time needed to make complex corrections is the same as the time taken to digitise a document from scratch. This is likely an over-estimate, as our testing has always found Extract's outputs to be at least partially accurate. However more data is needed to estimate the time savings in these cases.

We find the following distribution for the total time saving (%), suggesting a potential time saving of 40 - 55 %.

Confidence intervals

We used bootstrapping methods to calculate confidence intervals for our performance metrics and document number estimates. See scipy.stats.bootstrap.

All confidence intervals quotes are 90% confidence intervals.