

Arose Finance Specification

Data

Further to our conversation, please find attached:

- A subset of 50 Lender profiles in the 'Research Matrix' tab;
- 3 client profiles, I have manually produced this evening;
- 3 lender transcripts (running through their appetite & criteria)

We continue to build out the data set. I can generate more client profiles manually as/when appropriate – though the exercise this evening does make apparent that we don't have volume of readily available client profiles (because we do not have the profiling automated).

Crucially, we do not have client transcripts currently which we would utilise to automate the client profiling. I am currently using [8x8 Work](#) and [Otter.AI](#) but cannot find a method to connect the two to automate transcripts for our use.

Within the Research Matrix there are 5 categories of variables:

1. Lending Coverage
2. Borrower Type
3. Borrowing Type
4. Security Type
5. Product Details, distinguished as:
 1. Regulated residential bridging
 2. Unregulated residential bridging
 3. Refurbishment bridging
 4. HNW residential finance
 5. Revolving facilities
 6. Regulated development finance
 7. Unregulated development finance
 8. Development exit
 9. Commercial owner occupier
 10. Commercial investment

Some variables are determinants for the matching algorithm (like LTV, min/max loan size, 1st/2nd charge), others are purely informational and help the client decide between viable matches (e.g. interest rates, arrangement fees, legal nuances).

Variables should be filtered in a considered fashion:

- Y/N: if client profile has 'Y' for this column, we should filter for 'Y'. If client is blank for this column, it should not be filtered on
- Number: fuzzy matching (as lenders can be lenient) with comprehension of what the field stands for. E.g. Minimum loan size should be greater than or equal to the client loan size. Maximum loan size should be less than or equal to the client loan size.
- Client A was a conversion from a Carehome to a residential, technically the lender does not lend on Carehomes, however they did in this instance because it had planning permission to be converted into residential. I suggest we would need to add more variables to the research matrix account for this or more lender transcripts.

Full data description:

- 211 lenders
- 301 criteria

Process

As I see it, in Phase 1 we are looking to build the following core functionality:

- Initial call transcript, form + documents utilised to pre-populate fact-find + inform fact-find question set
- Interpret client conversation from transcripts, documents and fact-find to filter lenders
- Emails sent to lenders containing bespoke information (+ documentation), based on lender information requirements
- Centralised response centre for lenders with ML feedback loop (enhancing question-sets, documentation requested, lender variables and matching logic)
- Summary email sent to client with indicative terms, following lender feedback

Sources of real-time feedback:

- Variables & lenders added directly to research matrix
- Client transcripts
- Lender transcripts
- Lender feedback

Real time feedback consequences:

- Addition of variables (number of columns)
- Addition to lenders (number of rows)
- Update to probability % for matching logic
- Update to question set + documentation requested

Scope of works:

S t	Workflow	Inputs	Output	Categori sation	Incre ment al	Co re Te	Es tim ate
--------	----------	--------	--------	--------------------	---------------------	----------------	------------------

e p					Business Value	ch nol ogy	d Bu ild Ti me
1	During initial call broker asks client the 'initial question-set'. Broker bridges client to fact-find	Initial Question Set Client call transcript	Form (bespoke by initial call) Follow up correspondence Follow up meeting invite Fact-find pre-populated Fact-find question-set enriched	Initial call transcript, form + documents pre-populate fact-find Iteratively improved 'question sets' (initial and fact-find) + documentation requested	10 mins per case		
2	After initial call, client is sent a form to populate. Form contains list of questions and requests documentation. Documentation is collected via form. Fact-find cannot proceed without client providing	Client Documents and Answers in Form	Fact-find pre-populated Fact-find question set enriched	Initial call transcript, form + documents pre-populate fact-find Iteratively improved 'question sets' (initial and fact-find) + documentation	2 hours per case		

				requeste d			
3	Ahead of fact-find, data collected from form and documentation are used to pre-populate fact-find (and update 'fact-find question-set'). During fact-find client is asked 'fact-find question-set'	Fact-fin d questio n set Fact-fin d form Client call transcri pt	Comple te data set	Initial call transcript, form + documents pre-populate fact-find Iteratively improved 'question sets' (initial and fact-find) + documentation requested	2 hours per case		

Transcription taken from 1) initial call & 2) fact-find. Sources are:

- [8x8 Work](#) telephone calls (placed on PC and mobile app)
- Teams calls
- Zoom calls
- Cellular mobile phone calls

Transcoding of transcript, documents and form into research matrix profile mapping

4	<p>Broker is presented with concise approval process, manual verification required that the transcoding process has interpreted the client transcripts correctly.</p> <ul style="list-style-type: none"> • Broker can approve the transcribed interpretation of the client profile • Broker can edit the client & property profile where the transcribed interpretation is inaccurate • Where the broker makes manual edits, the transcoding logic is educated on where it was mistaken and becomes better at interpreting the client profile over time 	Present ation of client profile Manual edit/veri fication from broker	Approv al of client profile	Understa nd client conversat ion and filter lenders according ly	None		
---	--	---	--------------------------------------	---	------	--	--

5	<p>Match the transcoded client profile to:</p> <ol style="list-style-type: none"> 1) Research matrix (using exact and fuzzy matching where appropriate) with binary outputs and; 2) Enrich the matching with unstructured lender criteria transcriptions to enhance the matching 	Transcoded client profile	List of viable lenders with a % probability of application success	Understand client conversation and filter lenders accordingly	4 hours per case		
6	<p>Matching algorithm self-educates:</p> <ul style="list-style-type: none"> • Where application is successful (loan completes), algorithm is reinforced by the success • Where application is unsuccessful (loan does not complete), algorithm is educated by the experience. Algorithm interprets where the loan fell down, logs the reasoning and updates the probability of loan success accordingly 			Understand client conversation and filter lenders accordingly	None		
7	<p>Functionality to trigger emails that are bespoke to each lender (bespoke by: documentation & content shared);</p> <ol style="list-style-type: none"> 1. Where previous loans fell down with this lender for the same client profile, relevant suggestions in 'question-sets' are provided to the broker to flag where the subject loan could fall down 2. Lenders to be able to click their email to respond with binary 'Yes', 'No' or 'More Questions' <ol style="list-style-type: none"> 1. If 'Yes', lender to attach their Decision In Principle, further documentation required for FMA and further information they would require 2. If 'No', lender to submit rationale why it is a 'No'. Option to upload voicenote why 3. If 'Maybe', lender to ask further questions which are received by broker. Answers are collected (& recorded) by Arose and transferred to 	List of viable lenders with a % probability of application success	Templated emails to each lender with feedback collection	Emails sent to lenders containing bespoke information (+ documentation) Iteratively improved 'question sets' (initial and fact-find) + documentation requested Centralised response centre for lenders	3 hours per case		

	<p>the lender. Repeat back to step 2)</p> <p>3. Following lender feedback, matching algorithm educates itself:</p> <ol style="list-style-type: none"> 1. Where 'No' algorithm understands reason provided by the lender and embeds it as additional variable to their enriched logic 2. Where 'Maybe' algorithm updates Question-Set for client profile to ask these questions upfront (including new documentation). Variable is added to the research matrix. Algorithm includes additional variable as part of it's matching criteria. Bespoke email template is updated to include the answer to this question for all future 'output' emails to lenders 3. Where 'Yes', algorithm is reinforced 4. Option for lender to provide feedback on how they want email formatted, incorporated going forward 			with feedback loop			
8	Presentation of feedback collected from lenders in templated email which can be sent to the client by clicking a button	Feedback from lenders	Email template to client	Email sent to client with terms	1 hour per case		

Total 12 hours, 10 mins

To that end, I think it's important we think about:

- Ensuring we do not spend resource building something that will shortly be made redundant by advanced technologies
- Presenting as a human to lenders and clients
- Need a visible overview of matching logic and learning logic, editable by management
- As feedback loop increases number of questions, we need to be thoughtful around the number of questions we are asking clients (and when we are asking them). Need ability to enhance workflow design (improved by A/B testing)

- Lender's do update their lending policy. Binary criteria and leniencies change over time – matching logic must learn quick where there is a change in policy

Questions we should be able to answer:

- How can we 'plug' in to rapidly advancing technologies to enhance our product as their offering improves?
- What is the material upgrade to the matching logic by utilising unstructured client and lender transcripts? The structured filtering can be achieved now for free in Excel
- What is the proposed path forward to probability scoring?