International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

From Data to Donors: A Framework for Using Everyday AI to Strengthen Blood Donation Services

Md Juman Hussan
University of New South Wales, Sydney, Australia

Abstract — Blood services collect a lot of data. They know who donates, when, how often, and what happens to every unit of blood after that. Most of this data still sits in spreadsheets and databases, doing very little beyond record-keeping. This paper looks at a simple, well-known framework for AI, the same one taught in introductory AI courses, and asks a plain question: what could a blood service actually do with it? Using Australian Red Cross Lifeblood as a case study, this paper walks through four uses of everyday AI: predicting which donors are about to stop donating, forecasting blood stock before shortages happen, supporting donor screening questions, and using generative AI to reach donors in their own language. Real figures from Lifeblood's published donor study, transplant program, and research investment reports are used throughout to ground the discussion in actual numbers rather than hypothetical ones. None of the ideas here need advanced or futuristic technology. They need clean data, a clear question, and a narrow tool built for one job.

Keywords— blood donation, artificial intelligence, supervised learning, donor retention, data quality, Australia.

I. INTRODUCTION

Every blood service runs on the same simple cycle. Someone donates, the blood gets tested and stored, and eventually it reaches a patient who needs it. Behind that cycle sits a huge amount of data: donor records, test results, stock levels, and donation history going back years. Most of that data is used once and then forgotten. It sits in a database, does its job for that one donation, and never gets looked at again.

This is a missed opportunity. The same kind of data that helps a real estate agent price a house, or helps a bank spot fraud, sits inside every blood service in the world. The tools needed to use it well are not new or exotic.

They are the same basic ideas taught in any introductory AI course: take some input, learn to predict an output, and use that prediction to make a better decision.

This paper applies that simple idea to blood donation. It uses Australian Red Cross Lifeblood as a real-world case, drawing on its published donor study, its transplant research program, and its own research investment reports. It also draws on a well-known public framework for explaining AI to non-technical audiences, developed by Andrew Ng, to keep the discussion grounded and free of hype. The goal is practical, not theoretical: to show where AI genuinely helps a blood service, and where it does not.

II. THE DONOR BASE: WHAT THE NUMBERS ALREADY SHOW

Lifeblood already shows what happens when data and research work well together. Its Australian Blood Donor Study, or ABDS, built a biobank from over 9,000 consenting donors. Table 1 summarises the key figures from that cohort.

Table 1. Key figures from the Australian Blood Donor Study cohort (n = 9,064 consenting donors)

Measure	Value
Agreed to data linkage with external health records	97.2%
Wanted to receive their own test results	99.8%
Average number of donations made before joining ABDS	42.7
Average participant age	51.7 years
Identified as female	54.4%
Donated both blood and plasma	61.5%
Reported eating more iron-rich foods after donating	17.0%
Born outside Australia	25.3%
Of European ancestry	89.5%

This tells us donors are willing to be part of something bigger than a single donation, as long as they trust the organisation

asking. But willingness is not the same as reach. Even with this scale of donor goodwill, almost 9 in 10 ABDS participants are of European ancestry. Recruiting a more diverse donor pool is still a work in progress rather than a solved problem, and this is exactly the kind of pattern a predictive model could help track across thousands of donor records at once, something no person can do by hand.

Donor data can also reveal things nobody set out to measure. Lifeblood's own analysis found that active blood donors are noticeably more likely to be up to date with routine vaccinations than non-donors, shown in Figure 1. This is not something the study was designed to find. It came out of simply linking two datasets that already existed, which is itself a small example of the value sitting untapped in donor records.

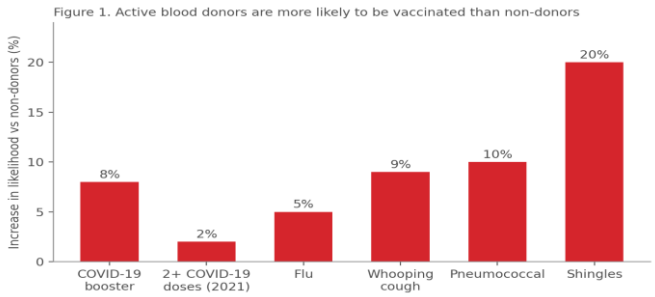


Figure 1. Increase in likelihood of vaccination among active blood donors compared with non-donors. Source: Australian Red Cross Lifeblood, 24-25 Research Annual Report.

III. A SIMPLE FRAMEWORK: WHAT AI ACTUALLY IS

It helps to separate three ideas that often get mixed together: Artificial Narrow Intelligence (ANI), Generative AI, and Artificial General Intelligence (AGI). ANI is AI built to do one thing well, like a smart speaker or a spam filter. Generative AI is newer and produces content such as text, images, or audio; tools like ChatGPT are an example. AGI is the far-off goal of a machine that can do anything a human can do, and it remains a distant research target, not something in use today. Table 2 sets these three ideas side by side.

Table 2. Three types of AI, and where each one actually shows up today

Type	What it means	Examples
ANI (Artificial Narrow Intelligence)	Built to do one job well	Smart speakers, self-driving cars, spam filters, web search
Generative AI	Produces new text, images, or audio	ChatGPT, image generators, AI writing tools
AGI (Artificial General Intelligence)	A goal: doing anything a human can do	Not yet built; a long-term research target

Almost everything useful in a blood service falls under ANI, not AGI. This matters because it sets expectations correctly. Nobody needs a general-purpose thinking machine to predict which donors are likely to lapse. They need one narrow tool trained for that single job.

Most of these narrow tools work the same simple way. You take an input, call it A, and train the system to predict an output, call it B. Table 3 gives a few common examples of this pattern, which is the same basic idea behind almost every supervised learning system in use today, including the large language models behind tools like ChatGPT, which are trained to predict the next word in a sentence, one word at a time.

Table 3. Examples of the input-to-output (A to B) pattern behind supervised learning

Input (A)	Output (B)	Application
Email	Spam? (yes/no)	Spam filtering
Audio	Text transcript	Speech recognition
English text	Chinese text	Machine translation
Donor age, donation history, time since last visit	Will donate again in next 90 days? (yes/no)	Donor retention prediction
Day of week, season, past donation volume	Predicted blood stock level next week	Blood stock forecasting

Data also comes in two broad types. Structured data lives in spreadsheets and tables: donor age, blood type, donation date. Unstructured data is things like images, audio, and free text: a donor's written comments, or a scanned health questionnaire. Blood services mostly work with structured data today, which is good news, because supervised learning methods handle structured data very well and do not need the huge compute power that image or language models require.

IV. APPLYING THE FRAMEWORK TO BLOOD DONATION

1. Predicting Donor Lapse and Retention.
A supervised learning model can take a donor's age, sex, donation frequency, time since last donation, and blood type as input, and predict whether they will donate again within a set window, say 90 days. Table 1 above already shows this kind of donation-history data exists and is already collected in a research-ready form, an average of 42.7 donations per participant before they even joined ABDS.

2. Forecasting Blood Stock.
The same basic technique, different columns, can help predict blood stock levels ahead of time. Day of the week, season, local events, and recent donation volume can all feed into a model that forecasts stock a week or two out. This does not remove the need for experienced staff. It gives them an early warning system that catches patterns across hundreds of donor centres at once.

3. Supporting Donor Screening and Questions.
Donor eligibility questionnaires generate a steady stream of text and yes-or-no answers. A narrow model trained on this kind of structured, historical data could flag unusual or inconsistent answers for a staff member to review, similar to how a spam filter flags a suspicious email. This keeps a human in charge of every real decision, while cutting down repetitive manual checking.

4. Generative AI for Multilingual Donor Engagement. Generative AI is a better fit for donor-facing communication than for clinical decisions. Lifeblood has already named multilingual consent forms and donor materials as a strategy for reaching migrant communities. Generative AI tools are well suited to drafting first versions of translated donor materials, reminder messages, or plain-language explanations of test results, always with human review before anything reaches a donor.

What Lifeblood's Own Research Already Shows

Measure	Value
Deceased donors worked up	545
Of these, proceeded to transplant	347
Kidney transplants	731
Liver transplants	194
Lung transplants	102
Heart transplants	85
New patients joining the transplant waiting list	1,405
Scientific workforce	158

This program is a more advanced version of the same idea driving donor retention prediction: use more detailed data, in this case genetic HLA data, to make a better-informed matching decision than manual review alone could produce.

Lifeblood's research investment gives a sense of where this kind of work is funded and prioritised. Figure 2 breaks down the 2024-25 research spend by program area.

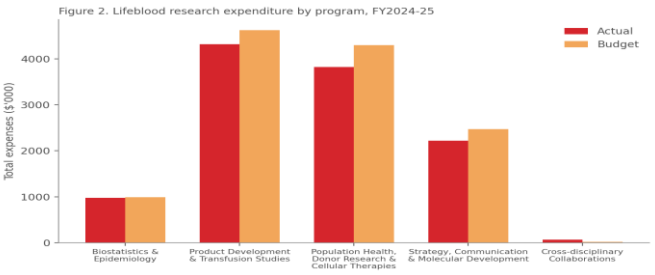


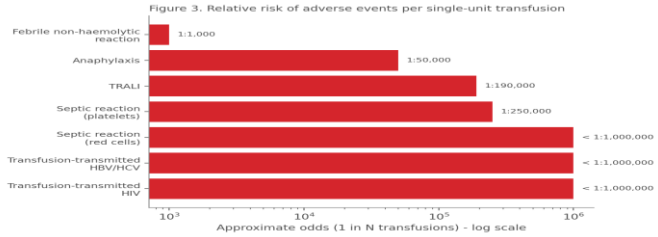
Figure 2. Lifeblood research expenditure by program, FY2024-25 (actual vs budget, $'000). Source: Australian Red Cross Lifeblood, 24-25 Research Annual Report.

Product Development and Transfusion Studies, and Population Health and Donor Research, together account for the majority of research spend, roughly 71% of the operational research budget between them. This is where donor-data-driven work, including the kind of predictive tools proposed in this paper, would most naturally sit.

Why Getting the Risk Numbers Right Matters
Any AI tool built on donor or transfusion data needs to be judged against a realistic baseline of harm, not an imagined one. Figure 3 shows Lifeblood's own published transfusion risk figures, which is a useful reminder that transfusion is already a very safe, well-studied process, and that any new prediction tool should be evaluated against how rare these events already are, not against a headline-driven sense of risk.

Figure 3. Relative risk of adverse events per single-unit blood transfusion, based on Australian statistics. Source: Australian Red Cross Lifeblood, Relative risk of transfusion (v4.0, 2019). This matters directly for AI applications. A model that flags donor screening answers for review, for example, will only be useful if it is judged against how rare serious transfusion-transmitted infection already is, less than 1 in a million for HIV,

HBV, and HCV. A tool that raises unnecessary false alarms could do more harm than good in a system this well controlled already.



Data Quality: The Real Bottleneck

None of these tools work if the underlying data is messy. Missing values, incorrect entries, and inconsistent record-keeping are common problems in any large organisation, and blood services are no exception. A donor record with a missing donation date, or an inconsistent blood type entry, will quietly damage a prediction model's accuracy without anyone noticing until the model starts giving bad advice.

There is also a well-known trap worth naming directly: collecting data for years before ever testing whether it is useful. The better approach is to start small, test a narrow model on the data already available, and let that early testing guide what new data is worth collecting next.

Limitations

This paper is a conceptual framework, not an empirical study. It does not report new statistical results, regression coefficients, or model accuracy scores, because no row-level donor dataset was available for this analysis. All figures, tables, and charts in this paper are built from statistics already published in Lifeblood's own study updates, posters, and annual reports, not from raw data accessed for this paper. Any future work testing these ideas would need a formal data-sharing agreement with a blood service and ethics approval, given the sensitivity of donor health data.

V. CONCLUSION

Blood services do not need a general-purpose thinking machine. They need a handful of narrow, well-built tools, each doing one clear job: predicting who might stop donating, forecasting stock before it runs low, catching inconsistent screening answers, and drafting donor communication in more languages. Lifeblood's own data, its donor cohort willing to share health records, its fast-moving research spend, and its transplant matching program already running on data at scale, makes it a strong candidate for testing tools like these. The starting point is not more data. It is a clear question, clean records, and a narrow tool built to answer that one question well.

REFERENCES

1. Ng A. AI for Everyone [online course transcript]. DeepLearning.AI; 2024.
2. Australian Red Cross Lifeblood. Essential to health: The Australian Blood Donor Study (ABDS). Research Study Update J25287. Melbourne: Australian Red Cross Lifeblood; 2025.
3. Dyer WB, Thomas K, Lindner G, Gemelli CN, O'Donovan J, Diaz P, Irving DO, Marks D, Karki S. The Australian Blood Donor Study and Biobank for Donor Health and Transfusion Safety Research [conference poster]. ABNA; 2024.
4. Australian Red Cross Lifeblood. Transplantation & Immunogenetics Services Research Program 2025-2026. Melbourne: Australian Red Cross Lifeblood; 2025.
5. Australian Red Cross Lifeblood. 24-25 Research Annual Report. Melbourne: Australian Red Cross Lifeblood; 2025.
6. Australian Red Cross Lifeblood. Relative risk of transfusion [patient information graphic]. Version 4.0. transfusion.com.au; 2019.