# Way forward for client identification across the health and social services sector

## Introduction and purpose

The Care and Delivery Task team is a task team of the Tanzania Enterprise Architecture Technical Working Group of the eHealth Steering Committee. The task team focuses on development of enterprise architecture and systems as they relate to client/patient level data.

This task team has had a series of meetings and workshops including meetings where issues of client identification across the health and social services sector have been discussed. The task team has also compiled information on the current status of client registration systems in various health programs, challenges in this area, as well as moves towards national legal identification and registration. This document is a proposed way forward on client identification across the health and social services sector. The document has been drafted by the task team for presentation to the broader enterprise architecture technical working group.

This document aims to expand further on ideas for a Master Client Index (MCI) / Client Registry as referenced in the national eHealth Strategy and national health sector Enterprise Architecture (EA).

## Background and context

### Overall context

Several health and social services programs in Tanzania are longitudinal in nature, in other words, they involve a series of encounters and services provided to a client or patient over a period of time. Examples include the continuum of care from ante-natal care through to labour and delivery, post-natal care and routine infant and under 5 care. Other examples include the longitudinal care involved in HIV care and treatment, TB treatment and diabetes treatment. Social welfare programs targeting vulnerable children are another example. There are also many interlinkages between such programs for example between community services and HIV care and treatment, between PMTCT in ante-natal care and HIV care and treatment, and between HIV care and treatment and TB treatment. In such contexts, it is important to be able to link data about the same patient/client over time and across vertical programs to ensure continuity of care.

In Tanzania there is a high degree of mobility between service provision points or health facilities. Patients/clients can receive services from community workers, their local dispensary or health centre and also attend a higher level hospital. They may also receive different types of services (advice and support, medical consultations, diagnostics/laboratory, and pharmacy services) at different geographical locations and from different service providers. In this context, it is important to be able to link patient data across service provision points for continuity of care. While resolving client identification challenges will not solve this problem in itself, it is a critical first step.

### Current health and social services identification systems

Many health facilities have a file system in which a client is given some form of ID number or file number when first seen at the facility. Each facility tends to have its own numbering system and these are not unique across facilities. Patients/clients often re-register if they have forgotten their previous file number at that facility.

In addition, various vertical health programs have ID numbering systems specific to that vertical program, such as the HIV care and treatment program and the TB program, whereas others have shorter term numbering systems such as ID number specific to one pregnancy rather than spanning all pregnancies of the same woman. These numbers are normally transferrable across facilities, as they are prefixed with a number referencing the facility which first issued the number, and patients are encouraged to use the same number if they receive services from that program in another facility. Volunteer services such as the HIV home based care program and the vulnerable children programs also issue ID numbers to clients prefixed with a number for the volunteer care provider. The facility or provider prefixes are also not consistent across vertical programs, for example the facility prefix number for the HIV care and treatment program is not the same as the facility prefix number for the TB program, and they are also linked to regional and district prefixes which become out of date when geographical administrative areas are divided. While the new facility IDs issued through the Health Facility Registry would solve the prefix issue, changing to these prefixes would have to account for legacy ID numbers issued in the past.

This results in patients/clients having multiple ID numbers. For example the same patient/client can have numbers issued by various facilities and numbers issued by various vertical programs. These identification systems assist with identification of a patient / client the next time they receive services from the same vertical program or the same facility or service provider but do not enable continuity of identification across the health and social services sector.

### Current legal identification context

The Registration, Insolvency and Trusteeship Agency (RITA) is responsible for registration of births and deaths in Tanzania. The birth registration rate is only 13.4% (National Population and Housing Census, 2012).



The National ID Authority (NIDA) was recently established [get date] and has issued IDs to [11 million?, also express as % of pop] Tanzanians with a number format of [put in format]. Meanwhile, the National Electoral Commission (NEC) did a voter registration exercise in 2015 resulting in registration of [23 million, also express as % of pop] Tanzanians with a number format of [ ]. The Prime Minister’s office, with assistance from the eGovernment Agency and the National Bureau of Statistics has the role of coordinating these bodies on the issues of legal identification, and chairs a High Level Coordinating Committee (HLCC) on Civil Registration and Vital Statistics (CRVS). Immediate steps being taken are that the Electoral Commission database is being harmonised with the NIDA database, and in future [timescale], NIDA will use the number format currently used by the NEC, to be called the Unified Lifetime Number (ULN).

NIDA relies on three sources of information

* RITA for proof of existance and date of birth
* PORALG for proof of current residential location of the person, and
* Ministry of Home Affairs for proof of citizenship or legal residence status in Tanzania.

The National ID is implemented under the Registration and Identification of persons Act of 1986 (R.E 2012), and currently applies only to persons aged 18 years and over, although this may be reviewed. NIDA has selected fingerprint biometrics as the primary method of checking for duplicate identity through AFIS (Automatic Fingerprints Identification System) Technology. The National ID Card is constructed using full polycarbonate material with an 80Kb micro controller chip, meaning that some data is printed on the card, but also additional data is stored on the mircochip on the card. NIDA are also open to storing more information in the chip if required by or useful to the health and social services sector.

## Rationale

### Why client identification ?

It is important to be able to consistently identify clients across time, across service provision points, across different types of services (diagnostics, consultation, pharmacy, community support etc) and across vertical health programs for several reasons.

First among these reasons is continuity of care. While identification of clients is not sufficient for continuity of care, it is a necessary prerequisite. Client identification greatly facilitates that the service or care provided at one point in time or at one service provision point is consistent with the service or care provided at another point in time or service provision point. This includes continuity of care within vertical programs and across program areas.

Secondly client identification is important for “the payer” ie health insurance schemes or Government or donor subsidised services, when clients do not pay all or any of the cost directly. The “payer” needs to verify whether the client was entitled to the service and to know the quantity and type of services provided in order to reimburse claims or justify funding and continue to provide that finance. Client identification can therefore contribute to sustainable health service financing.

Thirdly client identification is important for surveillance and monitoring, and quality of data. Without client identification, surveillance and monitoring systems encounter problems of double counting.

Fourthly client identification can assist with targeted communication between the health sector and the client to educate, inform and remind. When basic details are known about a client (geographic location, age, family linkages) this can make information campaigns more targeted.

### Why a master client index or client registry ?

Client identification does not simply refer to a unique ID number. Client identification involves using a range of information to establish whether this is the same person or a different person. Client identification is done using a set of demographic, biological and related attributes used to describe a client uniquely, for example name, date of birth, sex, face, geographic address.

Client identification in the health and social services sector is not the same as legal identification. Not every person who receives health and social services has legal identification, due to the fact that legal identification systems are not yet universal among the general population, due to the fact that some groups are not often legally identified (eg refugees) and also due to the fact that the health and social services system often sees an infant before the infant is registered legally. In addition some groups (eg MARPs) may not wish their legal identification to be linked with their health/social services identification, but the health and social services sector still needs to serve them both for their own benefit and for the benefit of the general population.

For these reasons, it is recommended that the health and social services sector has a client registry which is separate from (although linked to) legal identification systems.

## Way forward

### Governance/policy decisions

The following policy decisions need to be made (and may be revised over time):

* Whether or not client registry issues another ID number to clients (see appendix A) or uses existing client-held ID numbers
* What fields the client registry will store (see appendix B)
* What data quality standards and procedures need to be met by source systems before they are allowed write access to the client registry. Which fields are compulsory for source systems to supply and which are optional. There is a need to balance the advantages of opening up the client registry to maximum number of systems for comprehensive data vs the disadvantages of receiving very low quality data into the registry.
* Which systems or humans are allowed to create new registry records or update existing registry records, including which fields can be updated. Include discussion of roles and permissions of point of service systems, RITA and NIDA systems, insurance systems etc.
* Privacy policy and which humans or systems have access to read which client registry data
* SOPs for what the de-duplication and data cleaning processes will be, including matching algorithms (see appendix C) as well as human review of data quality and of possible duplicates

These policy decisions need to be documented in a living document which may change over time. The client registry software should allow adjustments of these policies over time. The governance structure for making policy changes will also need to be documented.

### Key features

This document is not a comprehensive requirements document; a comprehensive requirements document should be developed as part of a larger process. However in this document we suggest some key features needed in a client registry.

The client registry should have the following key features:

* The client registry should be a service which other electronic systems in the health and social services sector can connect to (authorised systems)
* Different authorised systems may be allowed to perform some or all actions in a configurable way (access control)
* The client registry should store data about a client which is useful for identification purposes. A client registry is not a health record, and data stored about the client should be limited to what is useful for identification and tracking/tracing purposes only. Possible data fields are discussed in appendix B.
* Authorised systems should be able to search for client records in the registry and access appropriate identification information about clients
* Authorised systems should be able to create new client records in the registry
* Authorised systems should be able to update changeable client information in the registry
* Authorised systems should be able to suggest merging of duplicate client records
* The client registry should have matching algorithms and features to establish probabilities of client records being duplicates or not, and allow the configuration of clear rules around automatic merging and human reviewed merging. Matching algorthims based on names and spellings need to be locally customised for common names and phonetics in Tanzania
* The client registry should have interfaces for human review of possible duplicates and of the quality and consistency of the data and interfaces for authorised humans to make changes as necessary
* The client registry should support configuration of policies within the registry, ie policies for authentication, privacy, authorization, integrity and non-repudiation (APAIN). The client registry should have clear audit trails.
* The client registry should be able to operate independently, but also to interface with and use Health Information Mediator services when they become available
* The client registry should be able to send and receive data with legal identification databases such as those at RITA and NIDA
* The client registry should support established data exchange standards and data standards

## Appendices

### Appendix A - considerations for ID numbers

The client registry may issue actual numbers to clients (a client registry ID) or alternatively the client registry ID number could remain internal, with clients identified through a combination of fields, including their other client-held IDs.

When designing IDs for making available to clients or other human beings, the following are considerations:

* Allocation large enough to cover full population over time
* Advantages of using serial numbers vs derived numbers from personal attributes or composite (serial and derived parts). Whether or not ID numbers reflect linkages between people (eg mother-child relationships, client-provider relationships)
* How IDs can be issued in a de-centralised way before the client is registered in the central registry (in absence of connectivity)
* Consider check digit to avoid human error in typing/writing (like the HFR which has xxxxxx-y)

How ID numbers can be stored

* + Written – handwrite, print out number on ID card
	+ barcode
	+ magnetic strip
	+ smartcard chip (NIDA has space on NID smartcard to include health data)
	+ store on client’s phone / sim card / device (issues with stolen phones)
	+ RFID and NFC
	+ Biometric such as fingerprint for MAT clinics (challenges with leprosy, children also)
	+ Eye Scans

### Appendix B - fields for consideration for inclusion

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Field** | **NIDA Field?** | **Definition and data rules** | **Discussion and data Source** | **Required (R) – Important(I) – Optional (O)** |
| Unified Lifetime Number (ULN) | Yes | This is a random number generated - with 12 digits which start with T for Tanzania. | RITA, NEC and NIDA are using.Endorsed GoT - SC meeting held June 2014.NEC registered people with fingerprints and gave ULN for Voter ID.NIDA cards that already have been issued, they will map and match these with NEC which uses ULN – they already have the NEC database.RITA was advised to start using these numbers for issuing new birth certificates, and people with old ones can go back and get re-issued with ULN. This will not happen until near the end of 2016. |  |
| Other non-health IDs including- Voter ID- Old National ID |  | Various formats documented in excel document |  |  |
| Birth Certificate Number | Yes |  |  |  |
| Health program IDs eg CTC ID, TB ID etc |  |  |  |  |
| Gender  | Yes | In Tanzania, there are two genders only. Coded list of choices, Male/Female |  | (R) |
| Client Name | Yes | This is the full name of the client, including First, Middle, Last. Name should be in three fieldsFirst\_\_\_\_Middle\_\_\_\_\_Last \_\_\_\_\_\_\_\_Other Names (Do we need this?) | How to address names with multiple spellings? Names can change over time. | (R) |
| Age and date of birth | No | Age is a number in years and months. Years \_ \_Months \_ \_ Date of birth is DD/MM/YYYY | Drop down of years?Months drop down 12 names?[Note: In most countries, use both to cover case where DOB is not known] | Either Age or DOB required? |
| Client Residential Address | Yes | This is the location where they live CouncilWardVillage/Mtaa | Drop down lists? There are issues with recording of residential location because of lack of lists of administrative areas (villages and wards) which are consistent, continually updated and linked-up over time, and also linking in with the post code system newly introduced by Tanzania Communications Regulatory Authority (TCRA).  | (R) |
| Name of head of Village /Mtaa  | No | Can this be linked if connect to drop down of address? |  |  |
| Mobile phone number | No | This is the mobile phone number associated with them. 4 digits for prefix\_\_\_\_7 for number\_\_\_\_\_\_\_ | A lot of people have mobile phones or have access to one; MNOs (Mobile operators) register millions of SIM cards, allow you to use Driver’s License, College University ID, letter from municipality, Voter Registration Card and/or Passport – this is what TCRA can use to validate people and who they really are at the MNOs. Issues with Tanzanians having multiple SIM cards, and often use family or friends phone instead of their own phone. GSMA have interest in Tanzania in having mobile numbers be used for IDs. | (R) |
| Place of birth (District) | Yes |  |  |  |
| Photo | Yes |  |  |  |
| Personal reference information | Yes | What does thiss mean? |  |  |
| Fingerprints L and R | Yes | Do we need? |  |  |
| Family linkages | No | Parent-child links, Guardian-child links, Spouse links |  |  |
| Other person linkages Treatment Support Person | No | Linkages with treatment supporter, community health worker  |  |  |
| Place where client encountered | Yes | Record where or who first registered client in registry, and also all subsequent places where client was seen  |  |  |
| Marital Status | No | Is this needed? |  |  |
| Education Status | No | Is this needed? |  |  |
| Religion | No | Is this needed? |  |  |
| Education Status | No | Is this needed? |  |  |
| Occupation | No | Is this needed? |  |  |

### Appendix C - matching algorithms

Some points to consider in putting in place matching algorithms

* Fuzzy matching eg DOB within three months, partial name match Alex Smith = Alexander Smith. Name matching algorithms and spelling conventions need to be customised for names in common local use. Phonetic transformation - language specific - eg “l” sub for “r” in Kiswahili
* Blocking - some minimum matching data in order to be even considered for comparison
* Matching algorithm determines Match / possible match / non-match – often done with weights to show which is most likely and ability to add more attributes for confirmation
* Probabilistic matching algorithm - must be customised country specific. Can learn over time with help of human review
* Use of human review to accept or reject matches suggested by the system
* Iterative use of more information (possible match, can we get more attribute values to confirm?)
* Tradeoff - Quality of data vs sophistication of matching algorithm (and time invested in testing and deploying matching algorithm)

####