

Opt-in, opt-out and bail-out

Privacy in new approaches to mental healthcare

Winfried Tilanus

Introduction

Right now a group of prominent mental healthcare professionals is seeking to reform Dutch mental healthcare. They call their project “De Nieuwe GGZ”, in English: “The New Mental Health Care”. This project has three aims: Clients and professionals should work together to get personalized diagnosis. This in contrast to the current situation where diagnoses are highly standardized. These standards are often the result of political processes, which results in clients receiving improper treatments and a vast waste of resources. The second aim of the project is to reform the current financial incentives in the Dutch mental healthcare system. Currently scientific tools that were never designed for N=1 measurements are used to determine the budget for individual treatments. On top of that, the focus on individual treatments in the financing system gives healthcare professionals very little incentive to cooperate or to seek for the best solution the whole system has to offer for an individual client. The last aim of this project is to reform the care process itself by improving the use of e-health. The first generations of e-health were translations of the traditional treatments to a digital environment. Though effective, the full potential of e-health is not capitalized by these digital equivalents of the traditional treatments. Modern ICT makes it not only possible to create more personalized interventions and to intervene wherever and whenever needed, but also to reform whole healthcare network. The ICT part of “De Nieuwe GGZ” explicitly aims to deliver an integrated and personalized experience based on new innovative interventions and the best interventions already available in the current healthcare network.

Though much of the project is still in its early design phase, some outlines of this 'digital transformation' in the mental healthcare sector are slowly becoming clear. The first objective of the ICT part is to create an integrated user experience. Right now every institution for mental health care creates its own e-health environment or portal for the clients. This results in a very fragmented user experience: logging in on multiple sites, having to enter the same information on multiple places and being bound to an ineffective intervention, because it is the only intervention a specific healthcare provider offers their clients. So the second objective, beside the integrated experience, is to be able to deliver 'the best of breed' interventions, regardless of which institution is the initiator of that intervention. Finally this part of the project aims to facilitate personalized interventions. When a client is followed over a period of time, some patterns that may help to optimize the interventions may become visible. For the personalization the project is also inspired by the notion of 'persuasive design' as used in marketing: a cluster of techniques that not only identify the potential needs of the customers, but also identify what incentives are most effective for influencing one particular user. Though this crossover between marketing and mental healthcare seems to be strange, it becomes totally logical when you bear in mind that both fields aim to change peoples behaviour. Finally the project wants to use Big Data techniques to investigate and improve the interventions. It should become a platform for ongoing innovation in mental healthcare.

In an early stage of the project, I was asked to take a look at the privacy implications of the ICT part of “De Nieuwe GGZ”. With this paper I want to open the public debate about it. Though there may be some legal and technological considerations in the background, the primary focus of this paper is ethical: Ethical considerations should precede legal and technical considerations, not the other way around.

Delivering an integrated experience

“De Nieuwe GGZ” explicitly approaches mental healthcare as a networked process. The client may interact face to face with a counsellor, may interact over the internet with a counsellor, may interact with a partly or fully automated site offering some kind of treatment or interact with wearable

devices like a smartphone, a smartwatch, a fitness tracker or a specialized sensor device. Each client will have his or her own mix of interactions. An integrated user experience means that all these interactions, from face to face counselling to internet of things interactions, are in line to deliver the best possible intervention to the client. The ICT part of “De Nieuwe GGZ” should follow the clients network and facilitate cooperation between the actors by facilitating data exchange between them. But each actor will have different data needs. When, for example, a smartwatch records the physical activity of a depressed client, then a counsellor needs an aggregated overview and an analysis of the movements. But an app on the smartphone that sends incentives to the client to get moving can adapt its interventions much better, when it has a real time feed of the current physical activity of the client. Each relation within the network will have its own data-needs, resulting in a network of specific data exchanges around each client.

The network may perform automated interventions. In the example of the smartphone app that uses measured physical activity to send well timed incentives to activate a depressed client, the smartphone app (or a third component) may calculate what intervention would be the best at a given moment. Such calculation can be based on fixed logic programmed into it, or it may perform statistical analysis or it may even use machine-learning techniques to optimize the interventions. Depending on the technique used, it may be more or less possible for humans to understand why the system does a certain intervention. When a fixed logic is used, its inner workings can be traced and analysed. When entering the realm of analysing big data sets or the realm of machine learning, it will take at least an expert to understand how the system came to a certain intervention, if that is possible to understand at all. Also when the interventions are entirely done by a counsellor or reviewed by a counsellor before they are performed, the counsellor still relies on the logic and the data analyses that is used to aggregate and interpret the data. The counsellor needs these overviews to deliver his or her part of the integrated experience for the client.

An other hotspot in the process of delivering the integrated user experience is the user interface. Not only will the user interface deliver the interventions of the different actors, because of the persuasive design, the user interface itself will become a way to intervene. It is the goal of the project to create an user interfaces that adapts itself to the learning style of the client. The way it presents itself and what kind of incentives are given depend on an analysis of the learning style of the client. In some cases it will be obvious that the user interface adapts itself, but in other cases it may be not so clear it does so, while it still may have a big influence on the interaction.

Delivering an integrated user experience to the client, results in a networked, rhizome like system of human and automated actors exchanging data to the need of their relations. Some of the actors within the network will react on feedback, some of them are adaptive in a more advanced way and some of them will be (automated) learning components. Such a network is a Complex Adaptive System.¹ The properties of such a system can't be understood by looking at its individual components, only by looking at the behaviour of the system as a total. Other important properties of a Complex Adaptive System, are a limited long range predictability and that the system as a total is hard to control.

Other data needs

Beside delivering an integrated user experience, the system has to meet some other demands. “De Nieuwe GGZ” is explicitly set up as a research project. So there should be some facility for researchers to access data generated in the network and for researchers to do their own measurements. The other data need is for performance measurement. Though the project wants to move away from protocolising treatments and move away from the use of tools that were never meant for N=1 measurements for decision making on individual clients, still some kind performance measurement is necessary to prove its effectiveness and to secure funding. The current proposal wants to do these measurements on an aggregated level.

Privacy and (mental)healthcare

Before taking a look at the general privacy aspects of e-health, it is good to clarify the notion of 'privacy' first. I define privacy as:

The ability to influence the process of mutual positioning within a relationship by controlling the information the other party has about you.²

This definition implies that privacy can only be understood within a relation. Sharing information with a third party equals creating new relations. An other implication is that information is seen as a way to influence a relation, to influence the 'mutual positioning within a relationship' to be precise. That positioning is a simultaneous process of 'being identified as' and 'to identify yourself with'. Privacy is seen as the ability to influence those processes by controlling information about you, it is explicitly bound to this positioning process. With that privacy is all about the balance of power within the relation. But it is also good to note that this definition is not normative: it does not say what balance of power is to be preferred. According to this definition privacy is also never binary: you can have more or less control over your information and with it, more or less influence on the positioning process within that relation.

As early as with the Hippocratic Oath³, the principle of not sharing information about the client with a third party has a prominent position in medical ethics. The main purpose of this medical secrecy, is to make sure a client feels himself or herself free to ask for help, regardless of what the problem is and regardless of any taboo or legal implication that is connected to the problem. This is not only a translation of the ethical principle that everybody should have access to medical care, it is also an import protection mechanism for the society: without this free access to medical care, people may stay untreated and become a danger for the society. Examples of this are the treatment of sexual transmitted diseases or the treatment of people suffering from paranoia and who may become aggressive over time.

In mental healthcare it is not only essential for the client to feel free to ask for help, for good mental healthcare, the client also must show his or her inner feelings and intimate thoughts. That makes a trusting relation between the client and the counsellor a necessity for a good treatment. Privacy is an important tool in creating that trust: knowing that whatever is told to the counsellor stays with the counsellor and does not have any repercussions outside the treatment, helps the client to gain trust in the counsellor. The counsellor uses that trust when doing interventions: the treatment can only succeed when the clients accepts the counsellors guidance towards behavioural change. To state it differently: mental healthcare can only work, when the relations are pure therapeutic and free from other disturbances.

“De Nieuwe GGZ”, new privacy challenges.

“De Nieuwe GGZ” approaches the care process as a network of relations. With that, the definition of privacy used here follows the structure of “De Nieuwe GGZ”: privacy is a property of each separate link within the network. “De Nieuwe GGZ” also has a strong normative aim: it wants to give the client control over his or her own care process. A high level of privacy for the client, a high level of control over the information the other party has about the client, is an important tool facilitate this control for the client.

The classic view on privacy dictates that the use of the data should be limited to its intended purpose, the client should be informed about what data is gathered and how it is used, the client should be asked for permission before processing the data, the client should be able to review his or her own data and the client should be able to withdraw the permission to process the data.⁴ For “De Nieuwe GGZ”, this would translate in (real time) insight in the data that is bound to each relation

within the system, granular choices in what to share or not and the possibility to determine how long the data is stored, (though that last choice may be limited because of legal demands). The medical secrecy would translate for “De Nieuwe GGZ” into strictly keeping all information within the relations to safeguard the client from interference from other parties. Other parties who have a legitimate interest in access to the data, like scientific researchers or financiers, should only receive aggregated data.

At first sight this strict privacy regime seems to resolve all privacy issues of “De Nieuwe GGZ”. Looking closer, several issues arise. First of all the granular insight and permission of “De Nieuwe GGZ” can easily result in dozens of privacy choices for one client. In practice people already tend to blindly click on 'agree' because they are asked too often to make privacy choices.⁵ This pattern is also visible in electronic systems for mental healthcare. Take for example “Jouw Omgeving”, an electronic environment for communication between clients with a light mental disability, their counsellor and their family. “Jouw Omgeving” presents a fine grained system for making privacy choices. I have observed that clients tend to allow everything or disallow everything, but they rarely make sensible choices. Having too much choice makes it impossible to make sensible choices, nor does it help when the client does not understand well what the consequences of the choices are. So a system of fine grained privacy controls doesn't necessarily result in the desired control for the clients of “De Nieuwe GGZ”, because it offers too many choices.

An other effect is that the persuasive design of “De Nieuwe GGZ” may create an 'e-health bubble'¹⁶. By adapting the information it presents, by adapting the interventions and by adapting the user interface, “De Nieuwe GGZ” may limit the choices the client has into a predefined direction. Once the client resides inside such a bubble, his or her choices become limited. Because it is not evident to the client he or she resides in such a bubble, and because it is hard to see what options there are outside the bubble, this effect also may severely limit the clients ability to make choices about his or her relations within the system.

“De Nieuwe GGZ”, as noted earlier, has the characteristics of a Complex Adaptive System. A Complex Adaptive Systems can't be understood by looking at its parts, only by looking at its functioning as a total. On top of that Complex Adaptive Systems have limited predictability. So controlling separate relations within the network can have unpredictable consequences for the interaction the client has with the system as a total. Because “De Nieuwe GGZ” has limited predictability, there is also a theoretical limit in the possibility to control it. This is not only limiting the possibility to influence the relations by controlling the information, it also rises issues about the safety of the system: it can't be guaranteed that the system stays within safe limits and that it doesn't encounter into sick-making interactions with the clients.

Finally it is good to notice that the clients of “De Nieuwe GGZ” are not always competent to make, sometimes complex, choices about their interaction with the system. A client that passes through a psychotic episode for example, is unlikely to be able to judge if a certain interaction is helpful or harmful. And because many clients of “De Nieuwe GGZ” are in vulnerable positions, “The Nieuwe GGZ” has an additional responsibility to guard their freedom of choice and to guard their safety.

Stepping up privacy for “De Nieuwe GGZ”

We have seen that in the context of “De Nieuwe GGZ” privacy is key to meet the targets of “De Nieuwe GGZ”, but at the same time privacy is problematic within “De Nieuwe GGZ”. We have seen that the classic way of dealing with privacy issues, giving the users direct control over their data, fails in delivering the privacy the project needs.

So, if giving insight and control over the clients data within the individual relations of “De Nieuwe GGZ” fails to deliver control and influence, should we then abandon it? One way to answer this

question, is from the ethical perspective. To understand the privacy impact of the Google Glass, Luciano Floridi⁷ introduced the distinction between diegetic information, information that is available within a given environment, and non-diegetic information, information that is only available outside a given environment. He uses 'transdiegetisation of the infosphere' for the development that information can be available inside or outside a given environment in a dynamic way. For mental healthcare the distinction between diegetic and non-diegetic information is very relevant: often there is a difference in availability of information between the environment of the counsellor and the clients environment. Over the years there has been a movement within mental healthcare towards making information diegetic in the environment of the client. This is for example reflected in the fact that the client has the right to see his or her own records. But in practice the relation is still very asymmetrical: in almost cases the counsellor can see the clients records in front of him with a press on the button, while the client has to go through cumbersome procedures to see his or her own record in a separate room. "De Nieuwe GGZ" takes as ethical starting point that the relation between the client and the counsellor should be symmetrical and that they should work together towards a personal diagnoses. This can only be done if there is also symmetry in the information position, if there is no difference between the client and the counsellor in what information is diegetic and non-diegetic. So the ethical starting point of "De Nieuwe GGZ" dictates that the client should have access to his or her own data, even if it is hard for the client to translate that access into sensible choices about the privacy.

There is also a practical approach to the question if we should abandon giving fine grained privacy controls to the client when such a fine grained controls hardly result in more control for the client. Luciano Floridi states that a big part of privacy comes down to user interface design. The way information and choices are presented can make a big difference in the ability for the clients to make sensible choices. Though there are limits to what level privacy controls result in control for the client within "De Nieuwe GGZ", there may still be room to optimize the clients control by optimizing the user interface.

So how would such a optimized user interface look like? First of all, when the paradigm is control for the client, then the client must consent to each relation (beside this being a legal obligation). So for every actor entering the clients network by participating in the data streams around the client, the client has to give permission. The client must, as good as possible, be able to understand where this permission is about, but also be able to make a fast decision. So asking for permission should be done in a short and easy to understand description containing who or what is going to participate, what data is shared and what will be done with that data. People change, their situations can change and their insights may develop. The clients may gain or loose trust. So it should be possible to reconsider choices previously made. A list of all requests for relations that are granted or denied should be accessible in a special screen of the user interface.

There is more relevant information then what fits in the short description, for example details on what information will be exchanged or how it will be used, but also more information about who the other party in the relation is and where to turn to when you want to know more or when you have a complaint about the relation. So each general description of a relation must have an option to view 'more information'. In there should be a more concise description of the actor, a possibility for contact when there is a problem and an example of the data that would be exchanged if the relation is not (yet) accepted or a full overview of the exchanged data for an established relation. If it is possible to exchange multiple kinds of data within one relation, each data stream must be presented separately.

Within a relation, different kinds of data may be exchanged and within a relation the data can be used in different ways. In such cases, once a relation is accepted, there are more choices to be made. Many of those choices follow naturally from the relation they are connected to. Other data streams

may be not so obvious from the relation or its description. And some other data streams or their usage may need explicit consent because of the consequences they (may) have. For each relation the choices of what data to exchange and what actions to take, should use sensible defaults. Choices that are obvious from the description should be on, choices that are less obvious, controversial or dangerous should be off by default. Of course the client should be able to diverge from the defaults. A logical place to access such choices would be from the 'more information' screen of each relation.

Finally the system may make choices that are not obvious at first sight or the system may make choices while it is not clear at first sight why it does so. Based on the historical behaviour of a client, for example, the system may adapt the timing and content of interventions. Or based on an analyses of the learning style of the client, it may adapt the interventions presented in the user interface. But an intervention may also be the result of a common decision with a counsellor. Making the choices in the system visible is an important step in empowering the client. So there should be an interface that shows what choices were made and what data and logic lead to it.

Bailing out: the ultimate control

Although an user interface as described above lets the client control as much as possible in an accessible way, we still can't expect an average user to actively use these privacy settings and we can't expect that the use of these settings results in control over the relations. And ultimately we will hit the limit that results from the system being a Complex Adaptive System. You can not understand the behaviour of the network when you are focussing on individual relations within the network and the network as a total has limited predictability. To state it boldly: privacy, defined as influencing the process of mutual positioning by controlling the data, is not possible when looking at the network as a total. Poking in individual relations does not result in control over the system as a total. There is even a severe theoretical limit to the controllability of the network: it can be influenced, but trying doing so may yield unexpected results.

So what kind of control can be given to the client? And how to guarantee patient safety? How to safeguard the system as total does not engage into interactions that are harmful for the client? Because of the unpredictability, control and safety can't be guaranteed on forehand, but we don't want to find ourselves concluding that in hindsight the system has been harmful or has interacted in a way that conflicted with the goals of the project. So the only option is to monitor the interactions as they happen and act when the interaction is not helpful for the client any more.

How to do such monitoring? How to detect derailed interactions? The person who knows best, is the client: the client is the first to experience the consequences of the interactions. Also the ethical starting point of "De Nieuwe GGZ", a symmetrical as possible relation with the client, dictates that the clients input should be leading here. At the same time the ability of the client to recognize and/or express his or her concerns may be limited. So a strategy offering multiple ways to monitor the interactions may be needed. A first strategy to get this information from the client is: giving the client an easy interface to express concerns about the current interactions, something like a "I feel bad about this" button. A second strategy is to actively ask the client: a short questionnaire should make it possible to get an indication about how the client experiences the interaction with the system.

These two ways of detecting derailed interactions are valuable, but they have important limits: they assume the client has a certain level of competence to recognize and express the experienced problems. On top of that, they rely on the interaction of the system with the client, while it is exactly that interaction we want to monitor. Their measurements are not independent from the variable they want to measure. So we also need detection strategies that don't rely on the interaction of the client with the system. To find such strategies we need to take a step back and ask ourselves: "how does it look like when a client is engaged in unwanted interactions with the system? Where

will that become visible?” One possible answer to that question, is that the client may become rebellious towards the system. The client may give fake answers to questionnaires or try to sabotage the system in other ways. Such 'deviant' behaviour may be a strong indication the interaction is not healthy any more. An other place where adverse interactions may become visible, is in the interaction with other people in the network: from counsellors to family members or friends, anyone in the network of the client may signal that the system as a whole is not performing as it should. Any of them should be able report such a signal, so it can be investigated.

How to act once a signal about unwanted interaction is discovered? A nice analogy is the situation a pilot of military aircraft faces when he or she observes unwanted behaviour of the aircraft. An aircraft is also a Complex Adaptive System that can potentially harm those involved. First the pilot has to check whether the observation is correct or whether it is a false signal. When the behaviour of the aircraft is really unwanted, for example because it is heading full speed for the ground, then the pilot can try to correct the behaviour of the aircraft. But when in the interaction pattern between the aircraft and the pilot harm to the pilot seems to be inevitable, for example when it is not possible to pull up anymore, then many military aircraft offer a safety-net: using the ejection seat to bail out. This makes it possible for the pilot to leave the potential harmful Complex Adaptive System before it really becomes harmful. That act of leaving the system is dangerous in itself: the structure of the aircraft may offer protection while the process of bailing out is a violent process of breaking of all bonds with the aircraft and taking distance from it, a process that usually involves explosives and rockets. And after bailing out, the pilot still has to land safely. He or she has to rely on an emergency parachute to avoid still being harmed while hitting the ground. So before bailing out, the pilot has to make an assessment: what is potentially more harmful: staying in this Complex Adaptive System that is going to harm me or bailing out with all possible consequences?

A similar procedure should be in place for “De Nieuwe GGZ”. Once signals are received that the system and the client are engaged in an unwanted interaction, those signals have to be investigated: are the signals correct? Can the interaction be adjusted so the interaction is again within its limits? Is it better to stay within the system or is it better to leave it? When a client for example, goes through a psychotic episode, the client may signal unwanted interactions and show deviant behavior. Though such a situation is hard to asses, someone has to check together with the client whether staying within the system is still the best option. And when leaving the system, there has to be a parachute in the form an alternative treatment: the client has to land safely.

When designing a bail-out procedure for “De Nieuwe GGZ” some elements must be taken into account. First of all, everybody involved must be aware that a 'maybe harmful' signal is an important signal that always needs a follow up. Putting such a signal aside is potential harmful for the client and is conflicting with the values of the project. While investigating such signals, the view should be from outside the system. The first assessment should be done by the client, together with a counselor and without any of the information produced in the relations within the system. Only when they decide it may be good to investigate if it is possible to change the behavior of the system, then they should take a look at the information created by the relations within the system. This outside view is needed to avoid the assessment of the 'maybe harmful' signals getting tainted by the potentially harmful interactions that need to be assessed.

Though bailing out is a severe measure, the option to bail out gives the client protection and an alternative when the system engages in adverse interactions. It functions as a last line of defense when the traditional privacy controls fail to meet their ends.

Conclusion: opt-in, opt-out and bail-out

From a privacy perspective, “De Nieuwe GGZ” is a challenging project. Its goal, a radical repositioning of the client into the center of the care process, implies strong privacy controls for the

client. At the same time the networked structure of the project makes these privacy controls less effective: it makes the behavior of the ICT-part of the project only predictable within small limits and that makes in its turn the system only limited controllable. The information can be controlled within the relations inside the network, but the effect controlling one relation has on the behavior of the network as a total is more or less unpredictable. A second tension between the goals of the project and its structure is that privacy choices become meaningless if the client has to make too many of them. Fine grained privacy controls for each relation within the network are infeasible if the client has to click away dozens of choices.

To overcome these problems, I propose a three layered system. The first layer is an opt-in for each relation within the system. The client must choose whether to engage into a relation within the system or not. Each relation comes with a sensible set of default privacy settings for that relation. The second layer is the possibility to further investigate each relation, to get insight in the information that is part of that relation and the possibility for fine-grained control over it. This layer gives the possibility to opt-out of any of the default choices from the first layer. This layer will probably be used by only a small part of the clients and it is questionable if using this layer will result in giving the clients really more control over their position. But the kind of relation “De Nieuwe GGZ” wants to have with their clients dictates that this layer has to be present: if it is the goal to put the client in the center of the process, then the client should have access to all information about him or her. The final layer is a safety-net in the form of a bail-out for clients. This bail-out consists of active detecting potential problems, a procedure for investigating potential problems and an alternative treatment when the interactions between the client and “De Nieuwe GGZ” move outside their limits.

- 1 Achmed, E. et. al. An Overview of Complex Adaptive Systems, (2005, <http://arxiv.org/abs/nlin/0506059>)
- 2 Tilanus, Winfried, Identity processes: Why we put our most intimate details on the net and are still worried about our privacy, Lecture at: Hacking At Random 2009 (august 18th 2009, http://www.tilanus.com/Identity_processes_proceedings.pdf)
- 3 https://en.wikipedia.org/wiki/Hippocratic_Oath
- 4 <https://www.rijksoverheid.nl/onderwerpen/persoonsgegevens/inhoud/bescherming-persoonsgegevens/>
- 5 Fox News, 7,500 Online Shoppers Unknowingly Sold Their Souls, <http://www.foxnews.com/tech/2010/04/15/online-shoppers-unknowingly-sold-souls/>
- 6 Tilanus, Winfried, Big Data, the e-health bubble and its fix, Presentation at TILTING perspectives 2015 (april 23th 2015, http://www.tilanus.com/resources/big_data_the_e-health_bubble_and_its_fix_concept_1.pdf)
- 7 Floridi, Luciano, Transdiegetic information: what it is and why it matters. (Keynote at TILTING perspectives 2015, paper not published yet, see <http://www.philosophyofinformation.net/articles>)