Submissions to the Human Rights and Technology Discussion Paper

I would like to strongly support the 'human rights by design' concept (Page 11, Proposal 13 of the discussion paper), as a way of capturing both the legal and ethics concerns associated with human rights.

I would further like to propose additional wording or working item/paper for developing a **generic computable policy framework** in support of the 'human rights by design' strategy, that can:
- support computer processable expression of rules defined by multiple legislative, ethics and professional code domains
- be informed by the latest international human rights developments and initiatives (such as GDPR) as stated in the discussion paper
- be tailorable to particular contexts of use, e.g. health, environmental protection, finance etc
- provide clear distinction of liability rules, including the expression of traceability and provenance between humans involved in the design and deployment of AI systems, and thus precisely support reasoning about accountable AI

Such a computable policy framework can be a part of the regulatory sandbox, where it can be implemented, tested, refined and deployed.

Ideally, such a framework should leverage or be informed by the existing relevant international standards (e.g. ISO, ITU-T, IEC, IEEE) that provide precise expression of rules over actions of actors in the system (and add more prescriptive form of regulation in addition to the principle-based regulation).

In fact, some considerations related to such a computable policy framework, can be found in a recently proposed foundational computational policy framework to support 'ethics by design' methodology, which also allows for the support of legal expressions related to the ethics and broadly, human right concepts. This framework was published in the paper titled "Ethics in Digital Health: a deontic accountability framework", attached in this proposal, and presented at the IEEE EDOC conference (https://ieeexplore.ieee.org/abstract/document/8945023.).

This framework can equally apply to 'human rights by design' concept, because of the generality of the concepts suitable for modelling both legal and ethics domains, the semantic foundation for which was ground in the deontic logic formalism. In fact, it is important to note that the framework is based on the semantics of the ISO/ITU-T/IEC ODP enterprise language standard and provides a rich set of accountability concepts, which are grounded in latest developments in deontic logic, as depicted in the figure below and described in detail in the attached paper. The framework can be used as part of a business/software methodology to provide a compliance link to the legislative and regulative rules associated with legislative and regulative structures. The paper shows how it can be link to a set of ethics principles such as do no harm, safety and reliability, explainability, data protection, contestability and compliance against policies such as privacy and consent.

Note that although the paper was motivated by digital health problems, it has general applicability, and we recently used it in the context of blockchain/distributed ledgers. Further, the

framework can be used to reason about the design as well as test implementation against the design and can distinguish accountability of parties involved in the design and deployment of AI systems, while allowing explicit labelling of actions of both human actors and automated decision makers, but treating them separately.

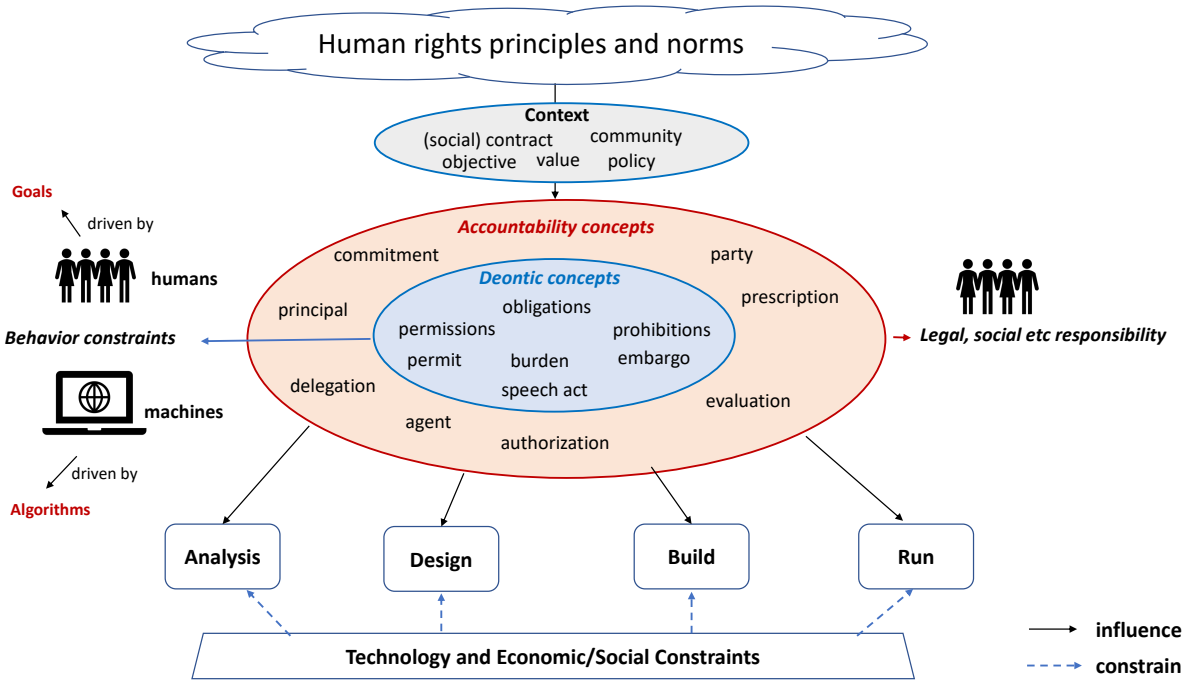

Figure 1-Deontic accountability concepts in support of human rights by design methodology

Dr Zoran Milosevic, FACS, SMIEEE
[redacted]

Ethics in Digital Health: a deontic accountability framework

Zoran Milosevic [1,2]
[1]Deontik, Australia
[2]Institute for Integrated and Intelligent Systems, Griffith University, Australia

Abstract — **We present key ethics concerns in digital health and introduce related ethics principles to address these concerns. We propose mappings of these principles into deontic logic concepts to support ethics-based analysis. This provides input into detailed design of deontic and accountability constraints using semantic foundation from the ODP enterprise language standard [1]. We refer to our approach as 'ethics by design' because it aims at explicit inclusion of ethics principles into contemporary software development and tooling. The paper is focused on digital health, but the approach has broader applicability.**

Keywords – ethics; digital health; privacy; consent; deontic logic; ODP enterprise language (ODP-EL); AI; FHIR.

I. INTRODUCTION

New digital health applications enabled by technologies such as big data, real-time streaming, advanced analytics and AI create many new opportunities for better healthcare. These technologies contribute to the collection and creation of vast amounts of patient information which can be used to provide safer and more efficient care and serve as a source of new insights in medicine. The sheer amount of data stored in multiple repositories and usage according to different policies, introduce risks and ethics challenges associated with inappropriate use of patient information. This includes inadequate support for privacy protection such as patient-defined consent policies. This protection should apply to primary use of data in support of a patient's care as well as secondary use, such as for research purposes. In the latter case, there is an additional ethics concern about the need to weigh up the potential benefits to society of using patient information for new research insights against the need to protect patient privacy.

Ethics challenges have become more prominent in recent times due to potential ethics issues associated with the use of AI. These include the impact of AI decisions on patient care, without full ability to interpret AI algorithm decision making process, or how to attribute accountability to such decisions in case of clinicians' use of it. There are increasing efforts to develop *ethics principles* to guide the design and implementation of AI enabled systems in general [3], and with specific digital health focus [4]. These frameworks provide an excellent starting point for a comprehensive analysis of ethics challenges in designing and building ethics-aware systems.

This paper builds on these efforts and proposes a structured approach for progressive refinement of ethics principles into formal models that can support reasoning about, designing, implementing and running AI-enabled digital health systems. We use a deontic-based formalism as a common theme for two aspects of the refinement problem. At the *analysis* level, we use it to facilitate the expression of the ethics principles to guide what the *system* should or should not do. At the *design* level, we use it for precise expression of behaviour constraints of *actors* involved in a digital health system, including their accountability. We use a model-based approach founded in deontic logic, drawing on our previous experience in implementing business contracts, while leveraging the concepts from the ISO/IEC Reference Model for Open Distributed Processing – Enterprise Language standard (ODP-EL) [1]. We believe that this standard offers excellent formal foundations for facilitating explicit inclusion of ethics principles in the design of digital health applications. We use the 'ethics by design' phrase to capture such ethics-aware design.

This paper assumes that the ethics principles are guided by particular moral values. The ethics questions of what social values are good and potential optimisation issues for conflict resolution, e.g. patients' moral questions mentioned before, are beyond its scope. The focus is on the applied ethics inquiry, concerned with a set of *norms* regarding what a person is obligated (or permitted) to do in a specific situation or a particular domain, typically drawing upon normative ethics such as deontological ethics [15]. This is broader than legal norms defined by a legal framework.

The next section presents emerging ethics concerns in digital health and identifies related ethics principles. Section III presents mapping from ethics principles to deontic statements to support ethics-aware requirement analysis. Section IV describes translation of such system level deontic statements into behavioural constraints that apply to the actions of the parties and system components to support ethics-aware design. Section V illustrates our approach with the privacy protection ethics principle. Related work and future directions are presented in sections VI and VII.

II. NEW ETHICS CONCERNS IN DIGITAL HEALTH

Big data technologies, advanced analytics and AI make it easier to process large amounts of data, make faster decisions and develop new insights. This is true both for the primary use of data in support of a patient's direct care, and for the secondary use, i.e. any application of data beyond the reason for which they were first collected [11]. Big data technologies make it easier to generate additional insights from secondary data through linking and aggregating multiple data sets, facilitating creation of new knowledge, but also raising many ethical concerns, which include, but are broader than legal concerns, as introduced next.

A. Privacy and consent - patient-controled data

Many current privacy concerns arise from large amounts of data, stored in different repositories and controlled by different organisations. In order to protect the private information that is collected, aggregated and processed for research purposes, many countries have created specific legislation or regulations. In Australia, for example, the Privacy Act 1988, provides extra protection over handling of personal and health information: organisations need an individual's *consent* before they can collect their health information. The Privacy Act recognises that it is often

impractical or impossible for researchers to obtain peoples' consent for the use of their data in specific research projects. As a result, a set of guidelines produced by the national data protection authority must be followed by any researcher approved to use health data without patient consent [11]. These guidelines also assist research ethics committees in deciding whether research projects should be approved.

Further challenge from using big data and advanced analytics over aggregate data can be *re-identification* of de-identified data, with the purpose of inferring some identity information from such aggregates and exploiting this information for say fraudulent purposes. In order to prevent such identity thefts, a number of privacy preservation techniques have been developed, such as differential privacy or secure multi-party computation [13].

Another challenge is requirements by many governments, in particular driven by the European General Data Protection Regulation (GDPR) policies, to provide greater rights for data subjects to control personal data, including access to it from external sources. This has implications on how one goes about providing a finer-grained consent policy that allows data subjects to provide access to their data for specific research purposes, while also keeping control over their sensitive personal health data.

B. AI concerns

AI technology brings additional ethics implications as discussed in [3][4]. They relate to AI making recommendations and augmenting activities of clinicians, e.g. image recognition, or medication dosage recommendations, while in future they may relate to more active functions, such as robotics applications. One concern is the so called 'explainability' problem [3], where clinicians do not fully understand the way the AI application makes clinical recommendations or decisions. Another concern is a possibility that the training data sources have bias towards some subjects which might cause the algorithm to behave unfairly. A further concern is to provide better trust to consumers through allowing them to challenge decisions or output of an AI algorithm, if it impacts them.

C. Ethics conflicts

As noted above, an important requirement is to allow an individual to define a consent policy to restrict access to (part of) their health information. On the other hand, there is another, conflicting requirement to do with the value of using an individuals' data to contribute to new medical knowledge and thus helping a new generation of patients. This may lead to an interesting question, "is it moral to benefit from research while opting out of electronic health records"? [11]. This was raised in the context of the Australian myHealth Record, where some 10% of Australians decided to opt out of this initiative for privacy reasons and for lack of trust in how their data will be used. This is a good example of a moral dilemma, along the lines of similar examples given in [14].

D. Regulatory and legal implications

It is now recognised that digital health, including AI, requires revisiting and augmenting existing regulatory and legal frameworks to address current variability and lack of standardisation in addressing ethics issues. This is needed both at the level of institutional review boards reviewing local research proposals and on a more global level, when governing ethical practice in multi-national, multi-institutional investigations [14]. The problem is exacerbated by the need to support transparency of use of personal information in a person-centric system, as for example driven by GDPR and HIPAA rules in the US [31]. This includes the ability for consent to be enacted or modified as part of data management issues, such as in support of data portability. The need for a better regulatory framework is also highlighted in the recent study [4], identifying 'a lack of clear rules, or even a tentative discussion framework, governing the ethical and social implications of patient data, AI and its growing use in the field of healthcare.'

E. Ethics principles

The ethics challenges presented can be addressed through agreeing on and applying a set of ethics principles. We identify the following ethics principles based on the recent proposals in [3] and [4], namely:

- Privacy data protection – must ensure that people's private data are protected and prevent breaches that could cause any damage to people
- Accountability – should identify people and organisations responsible for the design and implementation of digital health systems, including AI
- Compliance – must comply with relevant international, national, regulatory and legislative frameworks
- Safety and reliability – must ensure that systems are designed to avoid any negative impact to consumer
- Fairness – must ensure the training data for machine learning is free from bias that might cause the algorithm to behave unfairly against individual or groups.
- Explainability – must inform consumers about how exactly their data is used by an AI system and how it makes decisions
- Contestability – must allow consumers to challenge the output of the AI algorithm when it impacts them
- Do no harm – must not be designed to harm or deceive people through its decisions.

III. ETHICS-BASED ANALYSIS

Ethics is a branch of philosophy concerned with what is morally right or wrong. This paper's scope is on *applied ethics*, which is concerned with *norms* regarding what a person is obligated (or permitted) to do in a specific situation or a particular domain of action [15]. These norms or rules of conduct ensure that certain social values are maintained within the domains and protect the members of the domain from undesirable effects from the actions of others. Traditionally the members are people or organisations, but increasingly these can include automated systems, e.g. AI systems although the ultimate accountability rests with legal entities. In the case of AI systems, these are parties involved in the creation and operation of AI solutions. The applicability domain of norms varies, but the key point is that members of the domain must comply with the norms, either because they are part of the domain by choice or by inheritance, e.g. natural persons by birth in their country of origin.

A. Deontic logic underpinnings

Deontic logic is concerned with logic analysis of obligation, prohibition and permission (note that permission and prohibition can be described in terms of obligations [1]).

These concepts, sometimes referred to as deontic modalities, are a special kind of norms, and deontic logic can be a suitable formalism for logical analysis of applied ethics issues concerned with 'moral principles that govern a person's behaviour or the conducting of an activity'. Deontic logic does not look at the question of what kind of acts are good, or what is good. These are the concern of the meta-ethics branch of ethics [15]. Deontic concepts can be used to express ethics norms in various application domains, e.g. information privacy or biomedical ethics. Examples of norms in the information privacy are permission by an individual to define access to their information for a specific purpose and an obligation for a clinician or administrator to respect constraints defined by the individual.

B. Deontic constraints – digital health system

The ethics principles identified in section II.E express rules that specify the expected properties of a digital health system to satisfy ethics requirements. These rules can be treated as the deontic modalities that apply to the *system*, considered as an entity performing actions, including decision making that can affect consumers. These in turn can serve as an input to the detailed design and run-time enforcement which includes deontic constraints that apply to the *participants* involved in the system.

We begin with the *privacy protection* principle. This can be considered as an obligation of the system to respect privacy constraints for accessing personal health data, as specified by the consumers' consent (Figure 1). We take *obligation* to mean 'a prescription that a particular behaviour is required' as defined in the ODP-EL standard [1]. This high-level deontic constraint can then be refined into a number of fine-grained, design-level constraints, on actions of agents involved in controlling and accessing private health information. This includes individuals who specify consent rules (grantors), clinicians/researchers who access personal information for the primary/secondary use purpose (grantees), and authorities involved in the governance over the use of personal data, as will be elaborated in section V.

The *compliance principle* can be interpreted as an obligation of a system to respect the applicable regulation and legislative rules, such as the Privacy Act in Australia and related regulations to do with secondary use data. These rules will vary across different contexts, and some of the rules are illustrated in the example given in section V.

The *accountability* principle can be regarded as an obligation for a digital health system to identify *parties* legally responsible for the creation and deployment of the system. The ability to clearly represent chains of accountability and responsibility, including the links to appropriate legislative and regulatory authorities, increases consumers trust, in particular in AI enabled systems. We use the ODP-EL concept of *party* as it captures this intent. Party defined as 'a natural person, or any other entity considered to have some of the rights, powers and duties of a natural person [1]'. It is part of the ODP-EL accountability concepts described in section IV.C, but is introduced early in this section to facilitate remaining discussion.

Safety and reliability, has been a core principle of the development of medical technologies for quite some time [4], referring to the obligations of medical devices and clinical systems (i.e. their providers) to deliver services in a way that is unlikely to cause danger, risk, or injury to individuals. This is directly related to the *DoNoHarm* principle identified in [3] as an AI ethics principle, which states that 'civilian AI systems must not be designed to harm or deceive people and should be implemented in ways that minimise any negative outcomes'. This can be modelled as a prohibition of the system to create an algorithm that could cause harm to the application consumer. This prohibition should be traced back to the obligations of the system

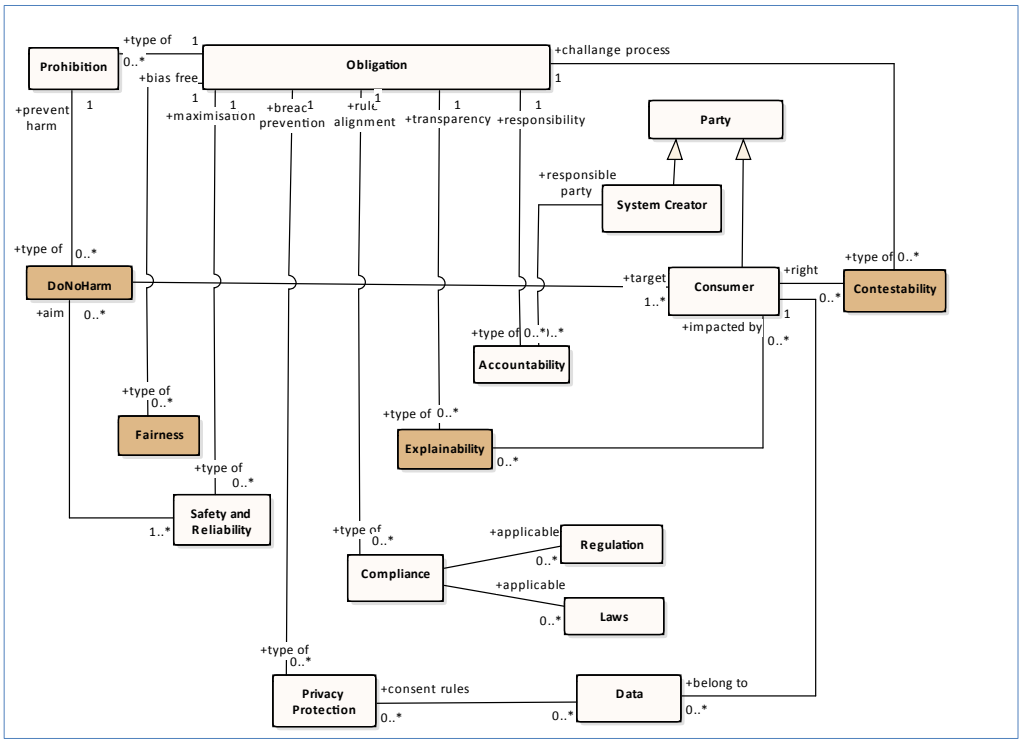

Figure 1 Ethics principles: high-level mappings to deontic constraints

creator not to design such a system, which is again manifested in their accountability, typically delegated to their organisation. Note that machine learning algorithms do not provide safety and reliability guarantees typical in safety critical systems such as pacemaker devices. This is because of their inherent stochastic nature, and further research is required to better position AI solutions in the context of such guarantees. A recent direction is in combining machine learning with automated reasoning techniques to build towards an explainable and dependable AI systems [26].

The *explainability* principle of an AI system is an obligation of an AI system to provide information to the users or consumers about how the AI algorithm makes a decision and which data set it is using to do so. This is of particular importance when such systems are used for clinical decision making to augment the work of clinicians. There are several techniques that assist in explaining AI's decision-making process, of which the LIME method (Local Interpretable Model-Agnostic Explanations) attracted a lot of attention recently[23]. Further, some authors propose the use of blockchain to track all the stages in AI algorithms as a way of understanding decision making processes. Such blockchain-based trails can assist to determine whether humans (and who specifically) or machines are at fault in case of accidents [20].

The *contestability* principle can be expressed as an obligation of an AI system to allow (i.e. give permission) consumers to challenge the use or output of the algorithm. This permission can also be considered as an authorisation given to the consumer to participate in the challenge process. Note that authorisation is an empowerment and is an accountability concept to be defined in section IV.C.

C. Accountability constraints – parties involved

Digital health systems include many types of party involved in the design, use and management of healthcare data and AI systems, with increasing requirements for clear statement of their legal, social or professional responsibility. There is a further requirement to be able to trace the way the rights and responsibilities of parties are linked with the system actions and their consequences. For example, individuals have permission conditions on consent associated with sensitive data fragments. This permission is passed to them through the act of a system operator, providing this additional behaviour capability. This action of the individual enables an automated system to act upon their personal data, for purposes such as analytics or AI algorithm processing.

A broader accountability framework is needed to describe a number of kinds of action that have different consequences on the future behaviour of the system; distinguishing these actions types provides framework for analysing the way responsibilities evolve [10]. These action types are defined as part of the ODP-EL accountability concepts to be introduced in section IV.C, leveraging the machinery of deontic logic. We thus distinguish between deontic concepts as general constraints on the behaviour of entities in the system, both IT systems and legal entities, and *accountability* concepts, which apply to the parties, capturing their legal or social responsibility. These separate but related concepts are useful in distinguishing between actions of human/legal entities and actions of IT components in AI systems. Parties can have intentions and are accountable for their actions as per the rules of the legal system. Many discussions of the responsibility of autonomous agents making their own decisions suggest a need to express chains of responsibility from one system to another, ultimately ending with the legal responsibility of humans or organisations involved.

IV. ETHICS-BASED DESIGN

The system level deontic statements of the form presented in the previous section can be translated into detailed behavioural constraints for the actions of the parties and system components involved. We use the key concepts from the ODP-EL standard to support precise specifications of such constraints, and position them in the context of the ethics principles, as described next.

A. Community concepts

The main structuring concept, called *community*, is used to describe a grouping of interested parties to achieve some

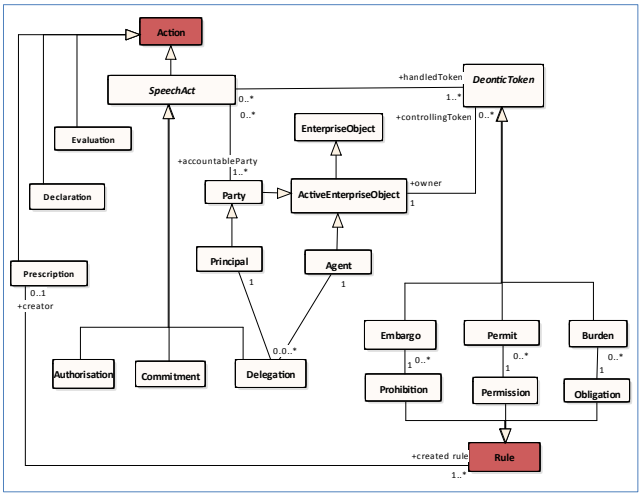

Figure 2: ODP EL: deontic and accountability concepts

objective, and their expected behaviour in support of that objective [17]. That behaviour is specified in terms of *community roles*, which can be filled by specific enterprise objects, provided their behaviour is compatible with the community roles. An *enterprise object* is any object in an enterprise specification, with special types of *party*, as introduced earlier, and *active enterprise object*, modelling an entity which can participate in some behaviour. The community concept was used to model various organisational structures in digital health [9], such as for electronic prescription, referrals, and a care team. Community roles can be fulfilled by humans, such as patients or clinicians, but also by IT systems. In the past, the IT systems were typically involved in actions in response to human actions, rather than making their own decisions. The increasing use of AI means these systems can also actively make decisions using their own algorithms and generating actions in response to these decisions. Thus, they can be modelled as active enterprise objects.

Recall that some ethics principles, in particular accountability, require that such AI-oriented roles need to explicitly support traceability to the parties involved in the creation of AI systems, to support the expression of legal responsibility. Several ODP-EL accountability concepts (to be introduced in section IV.C) can be used to express such

relationships formally. The accountability concepts can be also used to specify how certain community roles can be impacted by other roles, filled by active enterprise objects involved in automated decision making. This allows explicit support of the transparency, contestability or accountability in an AI-enabled digital health system. In such cases, AI systems can be modelled as a special kind of active enterprise objects, referred to as *agents*, with clear links to *principals*, representing legally responsible parties, e.g. those involved in the AI design or management. This feature also allows expressing compliance dependencies of the participants in relation to the legal, regulatory or organisational authorities, as required by the compliance principle in III.B.

B. Deontic concepts

The ODP-EL includes the modelling concepts of obligations, prohibitions and permissions. In addition, the standard provides concepts for modelling the dynamics of deontic constraints i.e. when they become applicable to the actions of parties and how they are passed among parties (and/or active enterprise objects). These are relevant for the governance, compliance and management of interactions between autonomous decision-making components and humans in a system. This is achieved by introducing a special type of enterprise object, called *deontic token*, which captures deontic assertions. The deontic tokens are held by the parties involved and holding one controls their behaviour [17]. Deontic tokens can be manipulated as objects while deontic constraints (e.g. obligation) cannot. There are three types of deontic tokens: *burden*, representing an obligation, *permit* representing permission and *embargo*, representing prohibition. In the case of a burden, an active enterprise object holding the burden must attempt to discharge it either directly by performing the specified behaviour, or indirectly by engaging some other object to take possession of the burden and performing the specified behaviour. In the case of permit, an object holding the permit is able to perform some specified piece of behaviour. In the case of embargo, the object holding the embargo is inhibited from performing the behaviour.

Another concept introduced to support modelling the dynamics of deontic constraints is *speech act*. This is a special kind of action used to modify the set of tokens held by an active enterprise object. The name was chosen by analogy to the linguistic concept of speech act, which refers to something expressed by an individual that not only presents information but performs an action [19]. Thus, a speech act intrinsically changes the state of the world in terms of the association of deontic tokens with active enterprise objects. This modelling feature fits well with the nature of AI enabled digital health applications, as it allows the specification of speech acts that can be performed by people and AI systems, yet distinguish them when necessary to establish links with ethics, legal and social norms. These modelling concepts are shown in Figure 2 (action and rule are the ODP foundational concepts [10] shown in red).

C. Accountability concepts

The deontic modelling framework presented provides a rich model to define many types of deontic constraints across for example AI systems and human actors. This framework is further extended to support traceability of obligations of parties, according to their broader responsibilities derived from ethical, social or legal norms. These extensions cover a set of accountability concepts to model such responsibilities [10], defined as [1]:

Principal is a party that has delegated something (e.g. authorisation or provision of service) to another. *Agent* is an active enterprise object that has been delegated something (e.g. authorisation, responsibility of provision of service) by, and acts for, a party.

Delegation is an action that assigns something (e.g. authorisation, responsibility of provision of service) to another object, e.g. agent. Delegation is one action type in ODP-EL related to accountability. There are other action types, that capture important business events in any organisational/social system and reflect the dynamics of communication among participants. They aim to precisely model how responsibilities evolve, which is important for logical reasoning about the ethical principles of accountability, data protection, and contestability.

Commitment, is an action resulting in an obligation by one or more participants in the act to comply with a rule or perform a contract. This effectively means that they will be assigned a burden. Examples are commitments by clinicians to deliver safe, reliable and effective healthcare to patients.

Declaration, is as an action by which an object makes facts known in its environment [10] and establishes a new state of affairs in it. This can for example be performed by an AI system (or a party managing it) informing the interested parties about the result of some analysis.

Evaluation, is an action that assesses the value of something. Value can be considered in terms of various variables e.g. importance, preference and usefulness. In digital health they can be various performance parameters used to express administrative performance, some accuracy or reliability measures associated with research findings or to assess the fairness of training data.

Prescription, is an action that establishes a rule. Prescriptions provide a flexible and powerful mechanism for changing the system's business rules at runtime, enabling its dynamic adaptation to respond to business changes [10]. This ability is important to support the applicability of new policies reflecting new legislations or recommendations arising from AI applications.

Authorisation, is an action indicating that a particular behaviour shall not be prevented. Unlike a permission, an authorisation is an empowerment. So, the enterprise object that has performed authorisation will issue a required permit and will itself undertake a burden to facilitate the behaviour. For example, the contestability ethics principle involves authorisation for the consumer to challenge AI decisions, giving them permit to do so by the AI system (or its creator/manager) who has the burden to enable it.

Figure 3 depicts the applicability of deontic and accountability constraints on automated and legal entities and their links with the development methodology, of which analysis and design were discussed above. A specific development environment would dictate a set of tools for build and run phases. UML-based model-driven tools with UML profile for ODP support can be a potential candidate [10], integrated with specific AI platforms. Further, specific technology platforms can be used, and they can impact the selection of controls, such as the some distributed ledgers, which provides new solutions for the implementation and protection of digital identifiers [18].

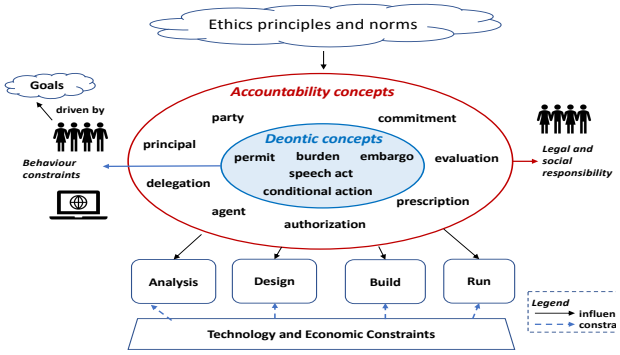

Figure 3: Ethics aware development methodology

V. EXAMPLE: HANDLING PRIVACY CONSENT

The privacy consent is essentially an organisational pattern that can apply to many situations and can be formalised as a 'consent community template' (Figure 4), a simple version of which includes the following roles:

- *grantor*, to be filled by any individual giving a consent, e.g. specifying permission for accessing ICU data but prohibition of accessing HIV data (if they exist)
- *grantee*, that can be fulfilled by professionals with the required credentials, i.e. clinician, to access grantor's individual health information for primary care purpose, or researcher for secondary use purpose.
- *consent authority* as a trusted enforcer, responsible for storing an individual's consent agreements, and monitoring the actions of the grantee role, to ensure that it acts according to the grantor's consent
- *national data protection authority*, responsible for defining and enforcing data protection policies
- *AI system role*, to be fulfilled by an active enterprise object that implements AI system functionality

The privacy consent community defines a number of deontic constraints that apply to the parties, including

- Obligation of the grantee to respect the privacy of the grantor, as defined by, say the consent authority role.
- Permission of the grantor given to the consent authority to store consent agreements, within a time period
- Permission of the grantor to the consent authority to monitor the grantee's access to the grantor's health information and enforce consent, in case of breaches;
- Obligation of the consent authority to reliably monitor access to grantor data, upon permission by the grantor
- Authorisation, of the grantor to the grantee to access the grantor's individual health information

The last of these constraints involves the grantor passing on a *permit* to the grantee to access the health record, and also places an obligation on the grantor (through passing the corresponding *burden*) to ensure this behaviour is enabled. For example, this obligation can require that security credentials are shared with the grantor. This *authorisation* action is also a *speech act* because it changes the deontic state of both the grantor and grantee. The effect of this speech act is that the existing grantor's permit to access its healthcare data is passed on to the grantee. For example, the consent directive gives permission to researchers to access the grantors ICU data but prohibits access to the grantor's HIV data. This illustrates the dynamics of deontic constraints throughout various actions associated with managing privacy consent.

The machinery of deontic tokens can be applied to define, implement, monitor and enforce other ethics aspects associated with the data protection principle. For example, many data protection rules defined by a national data protection authority set accountability and legal responsibility expectations for actions of researchers involved in using grantor's data. These data protection rules were established through the *prescription* actions performed by this authority, which essentially establishes obligations and permissions for all the parties involved in accessing patient data. This example also demonstrates how legal responsibility can be modelled using the ODP EL accountability concepts, where the researcher acting as a *principal*, with their own responsibility to the National Data Protection Authority, *delegates* some machine learning activities to an AI system, as an *agent* in this relationship.

The ODP-EL does not prescribe a language for the expression of deontic constraints, although the UML for ODP profile uses the Object Constraint language (OCL) [10]. One option is to use a concrete policy language presented earlier [7][8] for expressing the community, deontic and accountability concepts. The first element of the language is the concept of policy *context*, defined by the ODP concept of *community*, followed by:

- *policy activation* trigger, signifying that normative policies are in force; these can be temporal events or other events, such as violation of other policies,
- a community *role* to which modality and behavioural constraints apply (defined by the community context),
- deontic *modality* that applies to the party fulfilling a community role (subject role), which can also identify a target role referenced by the subject,
- event pattern specifying the expected behaviour of the party in terms of their actions and other occurrences, e.g. deadlines, or actions of other parties,
- violation conditions, specifying policies that can be triggered in response to the primary modality violation.

A *general policy* constraint is thus:

<communityContext><ActivationTrigger><role><modality>
<event_pattern><target_role><ifViolated>

Privacy consent policy, would then look like:

<ConsentContext><consentCreation>
<grantor><permission><accessPatientInfo>
<grantee>
<violation>

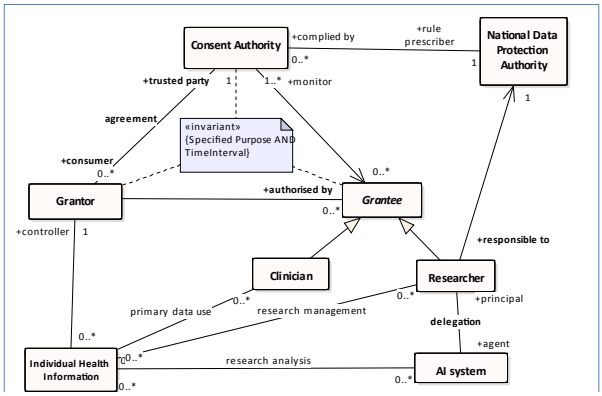

Figure 4: Privacy consent community model example

VI. RELATED WORK

The main aim of this paper is to develop a formal, yet implementable approach, for incorporating ethics principles

in the design and operations of digital health systems. We mentioned a number of related research and industry efforts in earlier sections but note the lack of systematic frameworks for o*verall ethics analysis and design*. An exception is an ethical design framework proposal for blockchain [25]. In terms of system design methodologies, the formalism which uses deontic, epistemic and action logic to reason about ethics, has similar aims to ours [21], but it falls short of implementation, stating that 'software emerging from this field is still in its infancy, no matter how impressive the theoretical background may be'. We adopt similar logic machinery but supported by model-driven software solution. More recent efforts by the same author [22], suggest need for 'design turn in applied ethics', to ensure that moral responsibility also covers design agents involved in the design of a system that end-users utilise. Our approach provides a sound foundation for expressing such responsibility, in part through the adoption of ODP-EL standard, and in part leveraging our previous work on business contracts compliance [7][8][12]. Other efforts focus on conceptual modelling of legal relations [28] and these can also provide input into our model if required to model the dyadic aspects of legal relations.

There are many recent popular articles about the ethical impact of AI, with an interesting discussion about significant uncertainty as to where AI responsibility lies – with parties who developed it, or parties who deployed it [27]. Our framework can accommodate both options.

VII. Conclusions and Future Work

Our approach is aimed at progressive refinement of ethics principles into software design artefacts. It adopts a pragmatic style of expression of deontic and accountability constraints suitable for distributed systems, including support for the dynamics of the creation and management of deontic constraints. This approach was demonstrated using consent resource of the recent HL7 FHIR standard [2].

Much remains to be done to develop explicit, ideally tool-based support of ethics principles into design of digital health systems. We believe our deontic accountability framework provides a novel contribution and we plan to continue this work, in alignment with the efforts presented in [3] and [4]. Our aim is to develop a toolkit for the 'ethics by design' approach while leveraging the semantics foundation of the ISO/IEC ODP-EL standard. This could involve the use of existing UML tooling, including support for UML for ODP standard [17], but also experimenting with broader set of tools and relevant formalisms for research purposes. For example, there is a need to support the concept of value, important for reasoning about ethical dilemmas and conflicts, while reflecting alternative behavioural paths similar to what is presented in [6]. The ODP-EL standard indicates potential direction for research in this area, using a rich model of *possible word semantics*, based on *Kripke model*, augmented with the concept of *utility* [1]. This model, grounded in legal compliance approaches [29] and broader ethics developments [30], can provide a firm basis for further tooling developments.

Acknowledgment

I would like to thank Dr Andy Bond and Dr Andrew Berry for their input to the earlier version of this paper.

References

[1] ISO/IEC 15414, *Information technology: Open distributed processing, Reference model – Enterprise Language*, 3rd ed, 2015.

[2] Hl7 FHIR, https://www.hl7.org/fhir

[3] Dawson D, Schleiger E, Horton J, McLaughlin J, Robinson C, Quezada G, Scowcroft J, Hajkowicz S (2019) *Artificial Intelligence: Australia's Ethics Framework*. Data61 CSIRO.

[4] Microsoft, *Healthcare, artificial intelligence, data and ethics* - A 2030 vision, Dec 2018.

[5] G.H. von Wright, *Deontic Logic*, Mind, Vol 60, pp. 1-15, 1951

[6] P. Linington, Z. Milosevic, K.Raymond, *Policies in communities: Extending the ODP enterprise viewpoint*, Proceedings Second International Enterprise Distributed Object Computing, 1998.

[7] Z. Milosevic, S. Sadiq, M. Orlowska: *Translating business contract into compliant business processes*, IEEE EDOC2006 Con.

[8] A Berry, Z Milosevic, *Extending choreography with business contract constraints,* IJCIS 14 (02n03), 131-179

[9] Z. Milosevic, A.Bond, *Digital health Interoperability frameworks: use of RM-ODP standards*, IEEE EDOC SoE4EE workshop, 2016.

[10] P.F. Linington, Z. Milosevic, A. Tanaka and A. Vallecillo, *Building Enterprise Systems with ODP, An Introduction to Open Distributed Processing*, Chapman & Hall/CRC Press, 2011.

[11] Australian Institute of Health and Welfare 2018. *Australia's health 2018. Australia's health series no. 16, Secondary Use of Health Information, section 2.5*, 2018.

[12] G. Governatori, Z. Milosevic, Dealing with contract violations: formalism and domain specific language, IEEE EDOC2015.

[13] D. Archer et all, *From Keys to Databases – Real-World Applications of Secure Multi-Party Computation*, https://eprint.iacr.org/2018/450.pdf, 2018.

[14] S. K Ludwin, T . Murray, *Dilemmas in medical ethics in the age of big data*, Multiple Sclerosis Journal, 2017, Vol. 23(10) 1306–1308

[15] https://plato.stanford.edu/entries/ethics-deontological/

[16] www.zdnet.com/article/is-it-moral-to-benefit-from-research-while-opting-out-of-electronic-health-records/

[17] P. Linington, H. Miyazaki, A. Vallecillo, *Obligations and Delegation in the ODP Enterprise Language*, the IEEE 16th International Enterprise Distributed Computing Workshops, 2012.

[18] G. Zyskind, O. Nathan, A. Pentland, Enigma: *Decentralized computation platform with guaranteed privacy*, 2015.

[19] https://en.wikipedia.org/wiki/Speech_act

[20] T. N. Dinh, My T. Thai, *AI and Blockchain: A Disruptive Integration*, IEEE Computer, vol. 51 no. 9, Sept 2018.

[21] J. Van den Hoven, G.J Lokhorst, *Deontic Logic and Computer - Supported Computer Ethics*, Metaphilosophy, Jan 2003

[22] Van den Hoven, J., Miller, S., & Pogge, T. (Eds.). (2017). *Designing in Ethics*. Cambridge: Cambridge University Pres

[23] M. T. Ribeiro, S. Singh, C. Guestrin *"Why Should I Trust You?" Explaining the Predictions of Any Classifier*, Proc. the 22nd ACM SIGKD Int. Conf. on Knowledge Discovery and Data Mining, 2016

[24] Dastani, M., Torroni, P., & Yorke-Smith, N. (2018). *Monitoring norms: A multi-disciplinary perspective.* The Knowledge Engineering Review, 33, E25. doi:10.1017/S0269888918000267

[25] Laponte, C, Fishbane, R., *The Blockchain Ethical Design Framework*, Georgetown University

[26] H. Bride, J. Dong, JS Dong, Z. Hou, *Towards Dependable and Explainable Machine Learning Using Automated Reasoning*. ICFEM 2018, pp 412-416.

[27] https://www.computerworld.com.au/article/661996/australian-businesses-split-over-where-ai-accountability-lies/

[28] C Griffo, JPA Almeida, G Guizzardi, Conceptual Modeling of Legal Relations, Int. Conf. on Conceptual Modeling, 2018, pp. 169-183.

[29] Governatori, G., *Thou shalt is not you will*. ICAIL 2015, pp 63-68.

[30] Moehler, M.,*The Rawls–Harsanyi Dispute: A Moral Point of View*, Pacific Philosphical Quarterly, Nov 2015.

[31] https://www.hipaajournal.com/hipaa-compliance-checklist/