Background

In this section, we will explore the concepts of digital identity and define key terms.

Identity is something we deal with every day. Every transaction we make with another entity relies on identity—recognizing with whom we are interacting. But however familiar we are with it, identity is still an extremely difficult subject to pin down. How do we know someone we know? If we have met them before and remember, we can recall what they look like, their mannerisms, the sound of their voice and their patterns of speech. Perhaps they are a friend of a friend, described well enough that we may recognize them on sight. Our family will have known us since we were born and can attest to our identity from birth. Our bank knows us by our bank account number and PIN, and hopefully gives us access to our assets. When we get pulled over by the highway patrol, we must present a driver's license, identifying ourselves as a legal driver. The highway patrol officer will also want to know our identity to know whether or not we have any outstanding warrants. All day, every day of our lives, we deal with our identities. We introduce ourselves to new friends, greet old ones. We announce our name when we call someone on the telephone. And every time we send an email, log on to a web site, or update our Facebook status, we are using our identity.

This thesis deals not with philosophical personal identity, but digital identity. Here, we will quote Phillip J. Windley's definition: "A digital identity contains data that uniquely describes a person or a thing...but also contains information about the subject's relationship to other entities." [5]

Identity can be defined as a set of attributes which make us individual, separate from others. In this sense, identity only matters in the presence of relationships between individuals. Without others,
there is no one from whom we need to distinguish ourselves. By being recognized, we are
differentiated from everyone else. Our identity is what makes us unique. State-of-the-art DNA
sequencing techniques allow for evaluation of our genomes. With this mapped, we can be identified
with certainty [1]. DNA is unique, that is, unless you are an identical twin. Fingerprints are unique,
even between identical twins [3], and are often used in criminal forensics to determine a person's
presence at a crime scene. By the laws of physics, your location in time and space is guaranteed to be
unique [4]. Services such as Google Latitude can use this information to track your identity, at least
as far as the movement of your cell phone.

But identity is also used in the sense that an individual can have multiple identities, each one
representing a subset of characteristics he or she wishes to present [10]. We refer to these as
attributes, specific pieces of information we reveal about ourselves. Some attributes may be
identifiers—strong identifiers, such as social security or credit card numbers, identify an individual
uniquely for a given population. Others attributes, such as height, hair color, and gender may be
weak identifiers in that they may not present a unique set [7][9]. Within the population of computer
science graduate students at Cal Poly, being a tall female with long brown hair is enough to make me
recognizable, but not within the university as a whole.

We may present out identity in such a way that we reveal only part of our set of characteristics. In
wishing to remain anonymous, we may present just enough to be useful without enough to actually
separate us from others. Or we may present a subset of our identity to one organization and a
different subset to another organization, thus creating pseudonyms. I have the account
"asbeug@gmail.com", which links my Google accounts, bank account, and utility accounts, but is
separate from my Facebook account, under "asbeug@yahoo.com". However, all accounts have
enough information to possibly be linked—my real name, address, and date of birth. The same
linking can be accomplished by all merchants with whom I use the same credit card. But when I use my "wanda.watinski@gmail.com" account, which does not have my name, date of birth, or address, I have effectively created a pseudonymous identity. I can use this identity across multiple transactions, linking them to each other, but not the rest of my transactions [5][7].

Identification is the act of using an identity, of binding it to an individual [5]. We may present our own identification or recognize the identification of another.

In stating our attributes as facts, we make claims. I have an attribute of being 32-years-old, but I may claim to be 43. The claims do not necessarily have to be truthful. To verify my claims, I can present credentials asserting the veracity of my claims. The iconic American credential is a driver's license. My driver's license states my date of birth, revealing my age at the time given and my authorization to purchase alcohol. Critical to all relationships is trust. Trust can be described as a belief in a separate entity, either a person or an organization, that the other entity will conduct a transaction as expected and desired. Trust assumes some level of risk in the context of the transaction [5][6]. When establishing an identity or presenting a credential, we are asking to be trusted—credentials are a method of transfer trust between parties. When I present my driver's license at the liquor store, multiple layers of trust must be established to complete the transaction of purchasing alcohol. First, the clerk must believe that I am the proper owner of the driver's license. She must also believe that my driver's license has not been tampered with in anyway and that the information printed on the license is valid. Finally, she must trust that the license was indeed issued by the State of California and have faith that the issuing organization has the procedures to competently verify my identity and produce a valid credential.

A digital identity management system (IdM) is a system of data, processes, and policies. The data contains information about identifiers and attributes for a given individual. IdM may also be viewed
through the different roles entities within the system play: subject, identity provider (IP), and relying parting / service provider (SP). A subject is an entity which the system attempts to identify. An identity provider is an entity which provides identities on behalf of subjects. And the service provider is an entity which relies on the identity provider to provide the identity to which it will provide services.

References