# Anonymity and IntelliDrive$^{SM}$

## Pre-Decisional Discussion Document

February 2009

Prepared for the Intelligent Transportation Systems Joint
Program Office, Research and Innovative Technology Administration

Contract #:  DTFH61-05-D-00002

For the best of reasons

# Anonymity and IntelliDrive$^{SM}$

## *Purpose*

The purpose of this paper is to define the implications of anonymity by design requirements on the IntelliDrive$^{SM}$ requirements and system architecture. The paper is intended to serve as a basis for discussion with stakeholders.

IntelliDrive requires both trusted communications between parties and the protection of personally identifiable information. The IntelliDrive architecture is currently being developed to assure anonymity and untraceability through the technical security architecture (as opposed to through regulation, policies, and procedures). Assuring that only legitimate, authenticated users can communicate and at the same time provide anonymity presents a complex challenge. As work has continued on the program, the negative impacts of this "anonymity by design" approach, in terms of complexity, deployability and utility have become clearer. This paper reviews the extent to which anonymous, untraceable communications by design is or is not required by the current *VII Privacy Policies Framework* and then discusses the complexity, deployability, and utility implications of attempting to provide such communications in the IntelliDrive program. Finally, a different approach is presented to illustrate that alternatives that preserve the privacy principles exist.

## *Background*

### Definitions

The terms "anonymity by design" and "anonymity by policy" have been frequently used in IntelliDrive privacy discussions, but they have not formally defined. In this paper, we propose the following definitions. "Anonymity by design" means that multiple technical controls have been built into the system to ensure that, to the maximum extent possible, a vehicle's or person's identity can not be determined based on their IntelliDrive data exchanges, or based on what was captured in one system's log file. This constraint would not apply to systems that by their very nature require individuals to opt-in voluntarily and identify themselves, such as toll payment applications.

Under the current envisioned implementation, one would have to tap into a real-time data stream and have access to logs from two separate protected certificate authorities in order to track an individual vehicle or obtain the sender's identity. A system that could enable vehicle tracking or identity discovery with access to fewer than two separate protected authorities would not qualify as providing anonymity by design under this definition.

"Anonymity by policy" means that a user's privacy is protected through adherence to written policies. If one person employed in the right place within an IntelliDrive system violated the policy, private data could be divulged. However, for IntelliDrive, many privacy protection mechanisms would still be included in the architecture. For example, personally traceable information would be stripped off at the earliest opportunity in the

**The IntelliDrive $^{SM}$ Logo is a service mark of the U.S. Department of Transportation**

data flow, and any information not needed for the specific purposes of IntelliDrive applications would not be retained. Policies, regulations, and possibly laws would be put in place to limit access to the data.

## The Privacy Principles

At the outset of the Vehicle Infrastructure Integration (VII) program, now called IntelliDrive$^{SM}$, there was considerable concern about whether such a program would lead to an unacceptable loss of privacy for drivers. The concern arose because a number of applications that were considered key to establishing the benefits of such a program involved communications between the vehicle and the infrastructure that could lead to vehicle drivers, owners, and passengers either being tracked along their driving path or being detected violating traffic regulations and law, possibly leading to legal action against them.

The National VII Coalition – which consisted of representatives from the US Department of Transportation (USDOT), automobile original equipment manufacturers (OEMs), and state representatives (including representatives of the American Association of State Highway and Transportation Officials [AASHTO]) – considered this issue and established a *VII Privacy Policies Framework*[1]. This document defines a set of nine privacy principles[2] that provide guidelines regarding privacy and personal information related to a national VII (IntelliDrive) system. The bulk of the framework would not be altered or affected by the issues discussed in this paper. Of the nine principles, it appears that three are particularly relevant to this discussion. These three principles (#1, #6, and #8) are cited below:

**1. Principle of Respect for Privacy and Personal Information** – *Commitment to respect for individual privacy in a National VII Program* [3]*means that VII-derived personal information should be acquired, retained, disclosed, and used only in ways that protect the privacy of individuals. Personal information users should collect, retain and use only anonymous information whenever possible. Users of VII-derived personal information and VII System administrators are expected to be accountable with regard to the personal information they collect and/or use in a National VII Program.*

**6. Information Protection and Retention Principle** – *Within a National VII Program, the VII System's technical architecture and structure should be designed to implement advanced security and other technologies to protect personal information against improper collection, disclosure or misuse in ways that may affect the privacy interests of personal information subjects.*

*Personal information users and information administrators should apply administrative,*

[1] The current version is the *VII Privacy Policies Framework, Version 1.0.2,* dated February 16, 2007.

[2] Please refer to the above cited document for a full statement and explanation of these principles.

[3] The term "National VII Program" is defined as: *the broad complex which, if deployed, would include all physical, technical and functional aspects of the subsystems and components used to collect, receive, transmit, store, and/or disseminate data and information, as well as the institutional structures and measures implemented in order to govern VII System users and administrators.*

*physical and technical controls appropriate to the protection of personal information derived from or obtained through the VII System. Particular attention should be given to:*

- *maintaining the security of personal information;*
- *protecting confidentiality of personal information against improper access; and*
- *assuring the quality and integrity of personal information collected or maintained.*

*Personal information users and information administrators should only retain personal information that is relevant to a valid purpose and only for as long as, and to the extent that, the information is protected against improper access, disclosure or use. Personal information users and information administrators should have data storage procedures that assure appropriate, secure disposal of personal information:*

- *when there is no longer a valid purpose for retaining the personal information, or*
- *when a stated or required time limit on data retention has been reached, or*
- *when data transmission has been completed within the VII System.*

*Identifiers, such as data addresses (potentially identifying a data source) captured during transmission or transport of data within the VII System should not be retained longer than is necessary to accomplish the data transport or transmission.*

**8. Participation Principle –** *In addition to receiving information regarding how personal information is collected and used in a National VII Program, each personal information subject should be expected to protect his or her own privacy. Personal information users should provide each personal information subject opportunities to:*

- *access personal information about himself or herself;*
- *correct any inaccurate personal information about the personal information subject;*
- *object to improper or unfair personal information use; and*
- *choose to remain anonymous, and not provide personal information.*

The first principle requires that *"personal information users should collect, retain and use only anonymous information whenever possible"*. This paper raises the question of whether or not the current anonymity by design approach is actually possible, when considering the cost and difficulties that have come to light.

The eighth principle requires that subjects (including drivers) may "*choose to remain anonymous, and not provide personal information."* For reasons discussed below, this paper raises the possibility that information on a subject's travel, although protected by regulation and policies, could be gathered. If this change where to be made, then the 6th principal becomes even more important. For example, under one revised approach, communications messages supporting traveler information and traffic management applications could, technically, be traced back to an individual vehicle, if captured in real-time. However once such messages are received and verified, there is no need to retain any information to support tracing the messages to an individual vehicle. Therefore, in accordance with the 6th principal, such information should be immediately stripped off and never stored.

An alternative to relaxing the eighth privacy principle would be to make it possible for subjects to "opt-out" of any IntelliDrive service.  Under the current concept, a subset of safety and public sector applications would be enabled at all times on all vehicles. This was considered acceptable because it provided the greatest benefits and the system would be designed to ensure anonymity. If anonymity cannot be guaranteed, then the ability to opt-out of the system entirely might be considered. This is discussed further in the *Anonymity and Participation Requirements* section.

## *Trusted Communications and Anonymity within the IntelliDrive System*

Safety applications conceptualized for the IntelliDrive System require mobile devices[4] to communicate – with the infrastructure or with other mobile devices or both – very quickly, with little delay in establishing the communication link ("low latency"). Also, the communications must be trusted – when a mobile device communicates with another mobile device or with an infrastructure device, the receiver of the messages must have confidence that the message is being sent by an authorized sender. This confidence is required primarily for safety applications and applications that involve electronic payments. The concept is the same as when a computer account user signs onto an account with some provider. Usually the account user provides an account identifier and a password known only to that user. However, the back and forth interaction required of this type of authorization is not anonymous, takes time, and may introduce a delay that diminishes the effectiveness of safety applications. It also requires an account "setup" process that may not exist between all entities that wish to communicate.

VII network designers chose to implement trusted communications through the use of anonymous security certificates[5] issued by a trusted Certificate Authority (CA) through a trusted, secure process. Once the CA has issued a security certificate, the certificate establishes the trustworthiness of the holder[6]. When that holder broadcasts a message, the message includes the certificate identifier, establishing it as a trusted message.

The use of anonymous certificates was the mechanism used in the "anonymity by design" approach that the IntelliDrive system designers planned to use. In this approach, mobile devices are issued "bundles" of anonymous certificates and can use any certificate in the bundle, for a pre-defined period, to establish trusted anonymous communication. This approach maximizes the anonymity of communications within the IntelliDrive network, since information that is traceable to a specific mobile device or user of that mobile

[4] Mobile devices include vehicles, but can also be other mobile items, such as person digital assistants (PDAs).

[5] A security certificate is encrypted data, stored in a computer, used to provide communications security. It can be used to verify that the sender/person is who they say they are or to encrypt data or both. Security certificates come in two flavors: *identity certificates* and *anonymous certificates*.

[6] If the certificate holder proves untrustworthy, the CA can revoke the certificate and notify others that messages accompanied by that certificate are not to be trusted.

device is never transmitted.[7] However, the use of bundles of anonymous certificates raises a set of complex technical issues.

### *Obtaining Anonymous Certificates*

First, mechanisms must be found for refreshing the bundle of anonymous certificates. Under the current approach, certificate bundles are given very short period (less than a week) before they expire[8]. If certificates were to be used for long periods of time, it would be possible to eventually correlate their use, and possibly compromise the anonymity of users. However, when all certificates expire, the entire bundle must be replaced.

If anonymity is to be preserved, this replenishment must itself occur in such a way that ensures the anonymous certificates can not be associated with a particular vehicle. If DSRC is used as the over-the-air communications medium, a large deployment of infrastructure to vehicle communications devices will be required. This is needed to ensure that all vehicles will be able to routinely access such communications, without requiring special trips. This large deployment "footprint," estimated to be on the order of 100,000 sites, would be required for the current security approach even if the initial applications only required vehicle to vehicle communications. Cellular networks or Wi-Fi connections could be used for certificate management if and only if adequate over-the-air link level encryption is used, in addition to end-to-end encryption. Furthermore, while cellular or Wi-Fi connections would be a possible alternative for certificate management, they can not be used for supporting IntelliDrive applications if anonymity by design is a requirement.[9]

### *Use of Non-anonymous Communications Networks*

The original IntelliDrive concept utilized the ITS Radio Service using Dedicated Short Range Communications (DSRC) in the 5.9 GHz band for all mobile communications. However, many IntelliDrive applications do not require use of this service. Leaving anonymity aside, cellular data services would be a feasible alternative for many non-safety applications. Cellular telephones can communicate securely – by encrypting the data and voice traffic that is being transmitted – but they don't communicate anonymously. One has to know which telephone is initiating a call and which telephone is receiving the call. Other wireless networks also focus more on protecting the content being moved on a network and on securing access to the network itself, rather than on

[7] In reality, even with the current anonymity by design approach, it is likely that if an entity had access to the records of both Certificate Authorities (e.g., through an appropriate legal process), the records could be correlated to allow at least partial tracing of an individual vehicle's communications.

[8] With short expiration periods for certificates, the risk always exists that a mobile device could need to communicate, for safety reasons, and not be able to do so because its certificate bundle has expired.

[9] Fisher, McGurrin, Hardesty, and Glassco, "Footprint Analysis for V2V Applications, Intersection Safety Applications, and Tolled Facilities," Noblis, March 2009, unpublished. The concern is that an unauthorized individual could, if the over-the-air link is not encrypted, establish when communications was occurring between an individual vehicle and the certificate authority, even though they could not read those messages. If this unauthorized individual could then gain information from within the identity certificate authority, he or she could use time stamps to determine which certificates were provided to that vehicle.

providing anonymity. The current anonymity by design requirement precludes the use of such common, commercial networks for IntelliDrive.

***Limited Utility of Probe Data***

Vehicle probe messages provide the essential information needed for traffic management, traveler information, and weather applications. There is no information in the probe message that identifies individual vehicles. However, even without such information, it may be possible to correlate messages with an individual vehicle, especially in low traffic situations. Moreover, through data mining techniques, it might be possible to correlate a large set of such probe messages from multiple RSEs, and associate them with a particular vehicle.  Access to a single source of information (the decoded probe message) would be adequate to identify individual vehicle behavior, violating the *anonymity by design* principle, as defined above.

The developers of the SAE DSRC message set standard have taken it as a design requirement that it be impossible to make such correlations. As a result, the standard requires the introduction of deliberate gaps in the probe records that each vehicle reports. However, it is difficult to introduce such gaps and at the same time retain the utility of the message for many applications. The current draft version appears to provide adequate information when roadside units are spaced far apart, as would be typical on freeways, but throws away too much information on signalized arterials, limiting the value of the data collected.

## Alternative to the Current Anonymity Approach

Privacy and the protection of personally identifiable information remains a critical concern for the IntelliDrive program. While the current “anonymity by design” approach is the one that provides the maximum anonymity consistent with the privacy policy framework, it is not *required* by the privacy policy framework. One could have mobile devices communicating with identity certificates, ones with longer periods of usage permitted, thus removing the issue of frequent updates to the certificate[10] and reducing the risk of a mobile device needing to communicate a safety message with an expired certificate. This would also allow the option of using other communications networks, such as cellular networks, whether for transmitting security information or for supporting applications that don’t require the high availability and low latency provided by DSRC at 5.9 GHz. Similarly, a relaxed requirement would reduce or eliminate the need for the data gaps that currently limits the utility of the probe information messages.

Using this approach would require additional emphasis on the Information Retention and Protection (6th) privacy principle. The information user must strip out the identifiers that uniquely identify an individual or device as soon as possible in the process of collecting and using the data being gathered.

[10] Security certificates normally have expiration dates. If the period for which they were valid were about a year, the certificates could be renewed when a vehicle owner takes the vehicle in for servicing. For non-vehicle mobile devices, similar mechanisms could be established for renewing certificates.

## Anonymity and Participation Requirements

IntelliDrive systems may be implemented either with *anonymity by design* or with *anonymity by policy*. In the latter case, privacy would remain an important principle, with personally identifiable information stripped off at the earliest opportunity. As suggested above, implementation of the system without anonymity by design is much simpler and less expensive, and allows the use of commercial wireless networks. The drawback, of course, is that anonymity is not guaranteed.

Participation in the IntelliDrive program may be mandatory or may be voluntary. In the former case, all new vehicles will be required to feature operational OBEs. Universal participation would be required in a subset of IntelliDrive applications because of the public good that would result, including safety and traffic management. In the latter case, drivers may choose not to buy new vehicles with an OBE or to turn it off and not participate in the program. The latter arrangement is sometime called "opt-out." (An alternate voluntary approach is "opt-in)."

| | **Participation Approach** | |
|---|---|---|
| **Anonymity Approach** | Mandatory participation<br>Anonymity by design | Optional participation<br>Anonymity by design |
| | Mandatory participation<br>No anonymity by design | Optional participation<br>No anonymity by design |

**Figure 1. Options for Anonymity Approach and Participation Requirements**

Conceptually any box in Figure 1 could represent the status of IntelliDrive applications as they come to wide-spread existence. The IntelliDrive community must determine which box will most closely define each of the applications. It is possible that the community will decree that the bottom left box, representing mandatory participation but no anonymity by design, will not be considered as an option. Similarly, the upper right box might not be feasible for some applications, e.g., those that require personally identifiable information in order to function.

### *Summary and Next Steps*

In summary, there is more than one approach for addressing the privacy concerns as stated in the privacy policies framework. The issues associated with the current "anonymity by design" approach are significant, raising significant barriers to a successful deployment. The next step is for the major stakeholders to determine whether or not the current anonymity by design approach is actually a program requirement. This discussion and resolution must occur quickly, and the implications affect critical aspects of the program, including not just the technical and security architecture, but viable deployment models and business models.