Features

SPIN provides the following features:

- **Distributed database** - prevents creation of a monolithic repository that is vulnerable to breach or misuse.
- **Distributed access controls** - engenders participation by the care delivery organizations and laboratories that provide data.
- **Query logs** - ensures accountability and oversight.
- **Anonymized clinical record analysis** - enables re-identification during a public health investigation. This process happens under institutional control, according to hospital policies and commensurate with the needs of public health and the certified authority of investigators.
- **Scalability** - extends participation, allows new data sources and applications, and facilitates voluntary collaborations in line with the goals of the National Health Information Network.

Distributed Database

SPIN leverages existing hospital databases and legacy information systems through a 3-step pipeline of extraction, transformation, and loading modules. First, patient records are extracted from local databases or extensible markup language (XML) files. Extracted records are anonymized, but retain sufficient information to detect clustering. Each patient record is assigned a random link identifier to support re-identification. Autocoding engines then transform text into a standard medical vocabulary. Finally, the anonymized and coded data are loaded into the peer database.

Institutions exchange digital certificates with approved peers to certify their identity and secure communications, forming *peer groups* that enable institutions to concurrently participate in multiple exchanges. Hub and spoke models are commonly used to minimize the number of peer relationships using a single entry point (supernode) for each group.

The SPIN-distributed query interface allows all members of a peer group to be contacted with a single query. Queries are performed by contacting the root supernode of the peer group, which propagates the message to each peer network or subnetwork until all peers are contacted. Results are aggregated asynchronously in reverse order. From the perspective of a client using the query interface, there is no difference between a SPIN network query and a local query.

Distributed Access Controls

The SPIN framework allows trusted agencies to certify the identities and roles of investigators. Each institution specifies what is allowed to be disclosed for each role according to that institution's policies through the use of a Distributed Access Control Framework.

Query Logs

Logging statements are recorded at each peer, and cannot be removed by external parties. These logs contain the certified identity of the investigator, the identity of the trusted agency who certified the investigation, and the time of query. Care providers are able to challenge the reasonableness of individual queries and deny access to agencies. Similarly, patients can audit care provider policies and investigator disclosures. This transparency is provided by the SPIN-distributed query, which returns the policies and query logs from all peers. Because all peer group members receive and log the same broadcasted query, a single institution cannot turn off logging or hide disclosures.

Anonymized Analysis

SPIN enables increasing levels of investigator access with peer-controlled disclosure.

Modern biosurveillance approaches rely on data mining to search for unusual patterns of disease. Hence, the algorithms require information on all encounters from all care provider locations. Using the SPIN approach, the Automated Epidemiologic Geotemporal Integrated Surveillance (AEGIS) biosurveillance system provides aberration detection, incurs minimal risk to patient privacy, and allows timely investigations to occur under emergency conditions.

Self-Scaling Architecture

SPIN architecture promotes individual participation and collaboration among members of each SPIN peer group. Autonomous peers form larger peer groups, and peer groups themselves can be linked to form larger, networked communities. Autonomy is central to this organizational trust, and ensures that care providers remain stewards of patient privacy.

The SPIN model leverages legacy information systems and existing institutional policies. For example, the federated identity and distributed access controls allow hospitals to continue using IRB and HIPAA authorizations. Other examples include the submission tools and query interfaces that extract, transform, and share data from existing databases using standard medical vocabularies.