While the broad industry shift to certified electronic health record (EHR) technology delivered legible digital records and more portable data, it came up short in quantifiably improving the cost, effectiveness and satisfaction of health care services. On average, it takes 17 years to integrate health care best practices into the flow of medicine and assessing care quality in time to make a difference with high-risk patients is simply not feasible due to fragmented and poorly standardized population health data sets.

Information technology boundaries are obstructing the delivery of effective and efficient healthcare. These boundaries make care more expensive and more difficult to navigate for everyone: patients, providers, caregivers, researchers and payers.

**Patients**
Currently, patients who receive care from more than one health care provider system are expected to be their own care navigators—suddenly tasked with managing critical data and processes across disconnected systems when they're sick and at their worst.

**Providers**
Doctors, nurses and other care providers are generally earnest people who want to help, but in reality, they have to spend valuable time focused on obtaining patient data from one disconnected, proprietary health record system to the next rather than caring for their patients and families.

**Caregivers**
As caregivers, family members and friends can provide an incredible layer of support but they frequently face overwhelming challenges helping their loved ones gain a foothold in their precipitous health journey.

**Researchers**
Researchers and population health scientists are eager to identify risk early and help patients before costs accrue and illness advances. Currently, the people conducting health care research face a disconnected landscape of data that requires time consuming cleaning before it can be used to find answers to a patient’s critical health care event—resulting in insight occurring months after it was needed to help the patient.

**Payers**
Payers want to implement value-based care models for health care like every other market has done—imagine buying a car and having no idea of the quality. Instead, payers grapple with claims and quality data that seem to have little to do with the actual clinical procedure.

Each of these stakeholders is negatively affected by legacy 1.0 EHR systems. In response to the Affordable Care Act (ACA) and Meaningful Use (MU) criteria, the health care industry rushed to purchase 1.0 systems before most technology standards were in place. The systems—built with proprietary data models at their core—resulted in long-term vendor locking and are primarily built on legacy programming technology from the 1980’s (see MUMPS). They rely on custom, one-way HL7 version 2.x pipes to connect data from producers to an evolving number of consumers. The technology is optimized for an era of client server architecture and does not translate well to the wonders of the cloud. These tools built to provide just enough information to submit the patient’s bill. They were never intended to effectively help patients on their personal health journey.

The current health IT ecosystem is constrained by its boundaries. There is a huge gap in how information technology functions in our personal life as opposed to our health care lives. We can move past the current state and overcome existing boundaries to deliver better, cheaper, safer health care for everyone.
The Perspecta HealthConcourse team and our coalition of partners are preparing for Health IT 2.0 solutions—ones that provide care without boundaries. Our unique approach is a 2.0 solution suite—an affordable, scalable and highly secure digital health platform (DHP) that embraces technology standards and makes care without boundaries possible. We built a plug-and-play platform based on standards and service orientation that leverages the best technologies available for the job. We focused less on specific point solutions, and more on cohesion across established solutions to provide a functional, harmonized ecosystem.

There are no proprietary data models at the core of the platform, which markedly reduces vendor-lock risk and the total cost of ownership over the product lifespan. HealthConcourse was developed to be future-focused—intentionally designed to minimize the cost to switch or add new capabilities when better or more cost-effective technology becomes available. As new requirements emerge and use cases surface, the HealthConcourse platform is configurable so new technologies can be integrated or the orchestration of technologies can be changed to easily address the problems at hand. To assure maximum flexibility, scalability and efficiency, we implemented the best open source and data processing technologies—NiFi and Kafka—at the core of the platform.

Leveraging the leading Kubernetes-based container platform in the market, OpenShift, our platform is portable, cloud native—yet agnostic to proprietary cloud technologies—and maximally scalable to meet the health care industry's evolving data needs. We also implemented a market-proven suite of process automation tools to cut away the many frivolous data tasks that keep health care providers from having time to talk to patients and work on their wellness plans.

The strength of our digital health platform, as illustrated in this paper, is its use of standards and open source technology rather than a dependency on proprietary data models. Our goal to build the world’s first “public health utility” can be best achieved through a community approach that values shared solutions and strives to continually improve and reduce costs for everyone.

1. Purpose and structure of this white paper

In this paper, we describe the architecture and use of our cloud-agnostic, fully standards compliant digital health platform (DHP).

We start by describing a technology-agnostic reference architecture for any robust digital health platform. We describe how a platform-based approach can bring order and structure to the digital health ecosystem where data is currently siloed and care processes are not shared across boundaries. We share this architecture openly to contribute to a collective, community-informed point of view on how to solve complex health IT, computability and interoperability problems with platforms and standards-based approaches. Implementations of this reference architecture are intended to solve health IT’s most vexing and costly problem—the lack of high quality and efficient cross-institutional patient care programs.

The second half of this paper provides a technical overview of HealthConcourse, Perspecta’s cutting-edge implementation of the DHP reference architecture. We outline technology and implementation choices and demonstrate how the reference architecture can be put into action. We offer details on a demo implementation leveraging a variety of technologies that mimics the complexity and heterogeneity of a real-world health care delivery IT environment. We close with recognition and attribution to our partners who share the vision of a multi-technology, best-of-breed, standards-based, service-oriented, platform-enabled ecosystem approach providing cohesion along with doors, ladders and bridges to overcome boundaries.

2. Solution components and reference architecture

APIs—the tip of the iceberg

Accessing and harnessing data across a variety of systems is not a new problem. For decades, engineers have approached this problem using messaging strategies, data file or database sharing strategies and interface strategies. Recently, the use of RESTful (Representational State Transfer) application programming interfaces (APIs), or RESTful web services, combined with the rapidly emerging FHIR (Fast Health Interoperability Resources) standard is accelerating and standardizing health care interoperability. FHIR-based, RESTful APIs enable health care data sharing across a federated and fragmented environment without necessitating data migration or locking up data in centralized solutions that sometimes reduces an organization’s control of the data.

However, the use of APIs does not, by itself, solve the problem of data fragmentation, lack of standardization or sub-optimization across a complex ecosystem. FHIR-based RESTful APIs are just the tip of the iceberg. To fully leverage the benefits of FHIR and APIs, more must be done—simply adding an API on top of an existing system of record is not enough—because the raw data is typically not in a state or format that lends itself to interoperability without manipulation. A recent review of data assets across the VHA concluded that the existing data “is unusable in its current state due to lack of normalization, duplication, being out of date and other noise and clutter. Our landscape analysis revealed a lack of a consistency when it comes to vocabulary, schema, and data formats.”[1] Even if these challenges are overcome, an API strategy ought to do more than wrangle data. Once data has been aggregated, standardized and made computable, new questions can be asked of the data and more can be done to automate routine data gathering and decision making processes. To optimize the health experience, a complete API strategy should:

1. Address the differences between the raw data from the systems of records (SOR) and the data requirements of FHIR and other interoperability standards. Data transformation and translation against industry standards helps to standardize the data into a canonical form
2. Hide the complexity of the underlying data environment to the data consumer. APIs and the platform behind them provide the opportunity to encapsulate the SOR and shield the data consumers from the underlying data silos and associated fragmentation. Aggregating the data from similar systems of record behind a single, more coarse-grained API reduces the burden on the systems of engagement that need this data. A thousand sources of data should not necessarily result in a thousand APIs.

3. Improve and enrich the raw data to maximize its use. Opportunities exist to enrich the data with provenance, security, privacy, tagging, conformance and other types of metadata.

4. Learn from the data. New insights in the form of calculations, inferences, assessments, recommendations, predictions, etc. can and should be derived from the data through analytics and evaluation of the data against evidence-based knowledge repositories.

5. Drive automation. Manual tasks are often time-consuming, error-prone and a cognitive burden on the end-user. Performing manual tasks requires a clinician to understand the best practice or clinical guidelines for a given health concern, to collect that data and assemble in their mind from a variety of sources, and perform calculations and evaluate logic that aid in clinical decision making. Much of this can be automated. When workflows are expressed in a computable, standard way, the execution of the sequence of activities, delivery of the data the activities require and or generate, and the evidence-based algorithms that support clinical decision making steps can and should be automated.

What is “under the API” is just as important as the use of APIs themselves. The APIs should be the gateway into a robust DHP capable of leveraging a variety of technologies to enable the APIs and maximize the usefulness of the data behind it. Further, a DHP and what is “under the API” is only as good as the orchestration used to provide cohesion allowing the technologies within a DHP to work together harmoniously.

**Orchestration—the glue that holds it all together**

A recipe without step-by-step instructions doesn’t make for a good meal. A shopping mall without a center aisle connecting the shops doesn’t make for a good shopping experience. A collection of health IT solutions without orchestration doesn’t result in a DHP.

Imagine that you have a risk calculator that can determine the likelihood that a patient has a certain medical condition if provided six different data points in a certain format encoded against pre-defined medical terminologies (e.g., LOINC). For the sake of example, two of the data points are numeric vital signs, one is a confirmed medical condition/diagnosis, one is a lifestyle factor (e.g., smoking or socioeconomic status), one is a family history of a medical condition/diagnosis and one is a diagnostic test result (e.g., imaging study or lab value). Now, assume that these six data points are captured and stored in different systems used by different providers working in different institutions. Further, assume that one of these six data points is a structured numeric value stored in an electronic medical record (EMR), the other vital sign is captured via a medical device—the confirmed medical condition/diagnosis is recorded as a SNOMED CT or ICD-10 code in a medical claim, the family history of a medical condition is reported verbally by the patient, the imaging study is available as a narrative report written in natural language (i.e., free text or unstructured data) and the lifestyle factor is inferred based on observation of patient behavior. In a digital world, the goal is to remove the burden from the clinician to:

- Understand what data is required for the risk calculator.
- Identify and extract the necessary data from the different sources.
- Transform the data from its native representation into the format required by the calculator—including extracting some data from free text.
- Infer or compute data where needed.
- Feed the data into the calculator so that an automated calculation can be performed.

The good news is that tools exist for all of these specific problems. Transformation engines can transform claims data into FHIR resources. Terminology engines can map between code sets (e.g., ICD-10 to SNOMED CT) or assign codes to textual descriptions. Natural language processing engines can extract key data from free text notes. Business rules, clinical decision support and other algorithms can be written to infer or compute data points. In today’s health IT environment, there is an abundance of point solutions that do a specific job very well. The problem isn’t a lack of technology, the problem is a lack of cohesion across all technologies so they can work together in the correct sequence, have their pre-conditions evaluated and confirmed and be coordinated with the appropriate handoffs. This is what orchestration does. It programmatically combines and sequences the technologies into a flow with clear lines of demarcation, clear handoffs and a clear end result. It makes the technologies work together harmoniously, even if they were created as independent units of software by different organizations in different programming languages at different times. Robust, configurable, scalable, governable orchestration is the single most important thing that a DHP must do well.

**Under the APIs—a reference architecture for a digital health platform**

This reference architecture describes the functions and components that constitute the solution “under the API.” The platform-based approach described by this reference architecture emphasizes modularity, interoperability and a common framework for optimizing how different solution components interact—are orchestrated—within a cohesive ecosystem. The reference architecture details the...
common elements required to effectively manage health care data across multiple sources of record in a fragmented and federated data landscape to reduce risk, cost and time to market. Implementations of this reference architecture, including HealthConcourse, provide the framework for delivering, managing, orchestrating and optimizing multiple health care services to address the needs of several stakeholders and use cases requiring access to the patient’s digital health record and other digital assets supporting health organization services (e.g., care delivery, customer service, operational management, analytics and asset management).

**Separation of concerns**

The DHP concept is predicated on the idea that health IT ecosystems comprised of multiple EMRs and other health-related data stores can no longer rely on a single “source of record” for managing the complete life cycle of health data. Further, no one system in a complex ecosystem can be solely responsible for managing data processing to drive all business processes or deliver integrated data views to all business users. The DHP reference architecture is an implementation of “separation of concerns”—a design principle for separating solutions into distinct layers, components or sections that individually satisfy a core business requirement or “concern.”

At the macro level, a DHP and the ecosystem it operates within are separated into three distinct layers as shown in Figure 1. The proliferation of multiple systems of records (where fragmented data resides) and multiple systems of engagement (where consolidated and standardized data is needed) highlights the business problem. Essentially, multiple systems of record exist, including EMR platforms, non-clinical business systems, data warehouses, patient-generated data stores, medical devices, applications and others that each contribute a portion of data to the overall patient record.

![Figure 1. DHP separation of concerns](image)

In today’s complex health delivery organizations, the systems of record are a mix of older and newer solutions developed in-house or delivered from a variety of health care solution vendors. In nearly all cases, the data across these systems of record are not aligned to a canonical form and have variable conformance to industry data models, data exchange and medical terminology standards.

When applying the “separation of concerns” principle, the systems of engagement are decoupled from the systems of record. We see this pattern applied in health care with health information exchanges, SMART on FHIR applications, mobile applications, patient portals, analytics dashboards and other systems of engagement where the end-user application is decoupled from the data sources.

The benefits of separating the systems of engagement from the systems of record include:

- Better modularity so that individual components can be changed or replaced without cascading impacts across the ecosystem
- The ability to have a best-of-breed environment with minimal vendor lock-in and greater portability of applications and components
- Longevity and portability of the overall architecture and ecosystem to ensure the systems can continue to provide business value as dependencies and infrastructure change over time
- Greater reusability of common capabilities to increase consistency, governance and time to market

Separating the systems of engagement from the systems of record does create challenges. The systems of engagement are encumbered by the fragmentation, lack of standardization and sub-optimization of data access and utility across a heterogeneous collection of systems of record. To address this challenge, the DHP fulfills the role of the system of insight.

A system of insight is an intermediary between the systems of engagement and the systems of record that hide the complexities of the underlying data architecture to ensure that systems of engagement have the information they need to satisfy business requirements. The system of insight is the solution “under the API.”
DHP core elements

Table 1 and table 2 define the core elements of a DHP. These core elements should be implemented as a set of modular services (or microservices) orchestrated to work cohesively as part of a well-managed ecosystem based on a service oriented architecture (SOA). However, each core element doesn’t necessarily correspond to a single unit of software or software component. The actual number, specificity and granularity of software components is open to interpretation provided that the following core elements are not collectively delivered as a single monolith application unable to achieve the benefits of extensibility, modularity, encapsulation and plug-and-play solution component migration over time.

After data is retrieved from the systems of record, a DHP can do a variety of things to make the data more usable and to then use the data for intelligent things before delivering it and other findings to the systems of engagement. Table 1 organizes DHP capabilities and services into four logical sub-divisions:

1. Data ingestion and standardization–responsible for standardizing and enriching data
2. Knowledge evaluation–responsible for managing and evaluating executable knowledge services
3. Business process automation–responsible for ingesting, running and automating executable process models
4. Secure data and knowledge delivery–responsible for delivering data and knowledge to consumers and managing appropriate access

Table 1. DHP capabilities and services

<table>
<thead>
<tr>
<th>DHP capability area</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data ingestion and standardization</strong></td>
<td>Standardization capabilities take data from the systems of record in whatever format available and standardize the structure and contents. This includes capabilities to map and transform inbound data structures to FHIR resources (or other preferred data exchange standards) and resolve differences in terminology and identities (e.g., patient, provider and organization identifiers). Data structure identifiers should be namespace to avoid identity collisions when data is aggregated from multiple sources. A key capability is the extraction of structured data from free text using natural language processing (NLP) to turn unstructured data into more discrete, structured data elements aligned to the canonical model. Enrichment capabilities enhance the base data resources by adding metadata, semantics and context to better index, describe, relate and/or classify the data. Common metadata includes data validation/conformance information, data classifications and information about data relationships. Data classification or segmentation is very powerful for sub-dividing the data to enable functionality like fine-grained consent, attribute-based access control and clinical research. Enrichment also includes capabilities that helps to manage the data relationships that exist once the data is aggregated. Managing data relationships includes identifying and reconciling data duplicates (e.g., duplicate FHIR resources) as well as creating linkages between records (e.g., relationships between different FHIR resources). For example, data reconciliation will weed out duplicate medications in a patient’s medication list, whereas record linkages will add information about logical, evidence-based relationships between medications and the medical conditions they treat.</td>
</tr>
<tr>
<td>Transformation services</td>
<td></td>
</tr>
<tr>
<td>Terminology services</td>
<td></td>
</tr>
<tr>
<td>Identity management services</td>
<td></td>
</tr>
<tr>
<td>Unstructured data and natural language processing services</td>
<td></td>
</tr>
<tr>
<td>Data validation services</td>
<td></td>
</tr>
<tr>
<td>Labeling and metadata services</td>
<td></td>
</tr>
<tr>
<td>Deduplication services</td>
<td></td>
</tr>
<tr>
<td>Record linking services</td>
<td></td>
</tr>
<tr>
<td><strong>Knowledge evaluation</strong></td>
<td>These capabilities are what differentiates a robust DHP from a standard message broker or service bus. Capabilities in this layer focus on management of and access to knowledge services (e.g., CDS, CQM and calculators) and analytical and statistical models (including artificial intelligence and machine learning models) to extrapolate insights from the raw data aggregated from the systems of record. This layer is the essence of a learning health system and is what allows the system to inform and guide clinical processes and decision making through evidence-based algorithms from vetted clinical research.</td>
</tr>
<tr>
<td>Real-time analytics services</td>
<td></td>
</tr>
<tr>
<td>Clinical decision support (CDS) services</td>
<td></td>
</tr>
<tr>
<td>Calculators and computation services</td>
<td></td>
</tr>
<tr>
<td>Clinical quality measures (CQM) services</td>
<td></td>
</tr>
</tbody>
</table>
The additional core elements in Table 2 are infrastructural in nature providing core and cross-cutting functionality. These elements could be implemented as microservices in a SOA or could be native capabilities of the underlying service management platform.

Table 2. Platform capabilities

<table>
<thead>
<tr>
<th>Platform capability or service</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business process automation</td>
<td>A data and knowledge platform is only useful if humans and other systems use it. Adding business process automation to a DHP provides a mechanism to incorporate, run and automate business processes to help contextualize where data and knowledge are needed. Models must be expressed in a computable format aligned to standards like BPM+. They should indicate where they need to bind data and knowledge to automate menial human tasks and optimize compliance to best practice and user experience. The models should be able to run as software in a run-time environment on top of process automation managers and/or workflow engines.</td>
</tr>
</tbody>
</table>
| Business process automation | • Computable, BPM+ models  
• Process automation managers  
• Model to data bindings  
• Model to knowledge bindings  
• Workflow and notification engines |
| Secure data and knowledge syndication | This layer includes capabilities to expose the aggregated, standardized and enriched data set and inferred insights to the systems of engagement. This is done through APIs based on standard web protocols (HTTP, REST) and health care standards (FHIR), and through asynchronous message exchanges such as “publish and subscribe”. Security controls for authentication, authorization and access control ensure appropriate access. Consent management appropriately filters data based on a variety of variables including the role of the data user, their purpose of use, their organization and other policy based filters. |
| Secure data and knowledge syndication | • Secure APIs  
• Access control services  
• Privacy and consent services |
| Data retrieval | SOR adapters retrieve data from the systems of record using messaging, API and database adapters so that data can be acquired and integrated into the DHP and subsequently transformed from the source format into a defined canonical model. |
| Orchestration | This area includes capabilities to manage the data processing pipeline by managing and executing configurable internal data processing workflows and calling business logic modules or external services as part of a SOA. |
| Persistence | This area includes capabilities for managing the data within the DHP as it is acquired, integrated and processed. This foundational capability includes the ability to store the data in long-term databases (e.g., data lake) or temporary caches and index the data for rapid search and retrieval. |
| Provenance | Data provenance is a critical DHP feature responsible for tracking all changes made to data structures and content as it passes through the DHP data ingestion pipeline. Every new resource or change to an existing resource results in the creation of a new provenance resource. |
| Auditing | These capabilities keep track of who accesses data and other pertinent details about system behavior. |
| Error management | Each contributing service or component participating in the DHP ecosystem should manage and raise exceptions in a common way. This capability provides a framework for consistently reporting and managing exceptions. |
| Surveillance | Surveillance monitors data into and out of HealthConcourse to trigger downstream actions and enable an event-driven architecture. Using the surveillance capability, solution components can trigger, or listen for and respond to, events based on the availability of new information. |
| Knowledge | Knowledge is overarching solution capability to represent, manage and invoke knowledge services (e.g., algorithms, business rules, decision support, analytics, etc.) on top of the data managed by the DHP. |
| Workflow | Workflow is the ability to define, manage and execute a sequence of activities relevant to a business problem, process or domain that runs within the DHP. |
| Automation | Automation refers to the ability to remove unnecessary human tasks and manual intervention in support of surveillance, knowledge execution, workflow or other DHP supported functions. Automation commonly occurs in the form of automated, contextually-correct data or knowledge delivery at the right time and place within a supported business process. |
Figure 2 provides a conceptual representation (reference architecture) of how the core elements and services—listed in table 1 and table 2—are organized, relative to each other, to create a cohesive system of insight. Any DHP implementation must be cloud aware (preferably cloud native in its design and architecture) and should leverage the computing, network, security and scalability capabilities of modern cloud platforms. However, to achieve a high degree of portability, the solution components of a DHP should be containerized into a common, robust container platform that insulates and abstracts the solution components from the proprietary features of any underlying infrastructure (e.g., specific cloud platforms).

This DHP reference architecture aligns well to health ecosystem descriptions from leading health care organizations such as Logica Health and Health Level 7 (HL7). The reference architecture incorporates several of the core services and capabilities described in the 2018 Logica Health Roadmap[3], including the use of the HL7 FHIR specification[4] for data transport and API definitions, Logica Health recognized standard medical terminologies, terminology services, the CDS Hooks specification[5] and data segmentation used for patient right of access and consent[6].

The reference architecture also incorporates several of the core services defined in the HL7/OMG Health Services Specification Project (HSSP)[7], including message transport services, terminology services, health services directory services, record locator services, identity management services, data retrieval services, data mapping and transformation services, event management services, health care security services and clinical decision support services. Implementations of this reference architecture are encouraged to be faithful implementations of the API specifications and related requirements for the services defined by FHIR, Logica Health and HSSP.

3. HealthConcourse—Perspecta’s DHP—enabling care without boundaries

HealthConcourse is Perspecta’s DHP based an open architecture and open, standards-based APIs built on the FHIR standard. HealthConcourse is a FHIR-first solution. FHIR resources are used as the canonical data model within HealthConcourse, and FHIR API specifications are used to define the primary set of HealthConcourse APIs. HealthConcourse transforms the raw data retrieved from the systems of record into FHIR resources. The data ingestion pipeline of HealthConcourse includes a configurable collection of services that implement many of the capabilities defined in the DHP reference architecture above (e.g., terminology services, NLP services, identity management services, etc.). Serialized FHIR resources are used as the input and output of most of the microservices, resulting in standards-based internal interoperability that enables modularity, encapsulation, agnosticism and a plug-and-play approach.

Figure 3 extends the DHP reference architecture by identifying the HealthConcourse technologies and solution components used to implement the services and capabilities illustrated in the reference architecture. Technologies in the “core platform” are integral to HealthConcourse and are part of any HealthConcourse implementation. The technologies and solution components in the services
The technologies actually implementing the capability can be replaced with alternatives. Elaboration of this technology plug-and-play aspect of HealthConcourse is further illustrated in figure 3.

**HealthConcourse core platform**

The HealthConcourse core platform is made up 5 core components.

- **NiFi and Kafka**: NiFi and Kafka are industry proven, open source technologies. HealthConcourse includes a very robust data orchestration architecture using NiFi and Kafka that handles all data retrieval, ingestion, orchestration, auditing, provenance and error handling. NiFi was originally created by and for the National Security Agency (NSA). It was architected from the ground up to address security, high data throughput and scalability. Likewise, Kafka—initially built by LinkedIn to handle massive amounts of data transactions and stream processing—is built with high data throughput and scalability in mind. Data is acquired and integrated into HealthConcourse through adapters managed by NiFi and/or Kafka (depending on the exchange patterns and data access mechanisms provided by the systems of record). HealthConcourse will actively pull data from systems of record or will deploy listeners that receive data passively from systems of record. Each system of record can provide a different data access mechanism and use a variety of data exchange patterns. As such, HealthConcourse leverages and extends NiFi's extensive collection of adapters and highly configurable, scriptable data ingestion capability and uses Kafka's library of adapters and internal asynchronous communications capabilities. The flexibility afforded by using these two tools allows HealthConcourse implementations to have custom orchestration flows so that each customer or use case can use the best data processing flows, controls and microservices for their requirements.

- **Smile CDR FHIR Server**: The Smile CDR FHIR Server is best-in-class. It is the hardened, enterprise-strength, commercialized version of the ubiquitous HAPI (HL7 API) open source FHIR Server. Smile CDR (and HAPI) is a complete implementation of the FHIR standard. Smile CDR is part of the core fabric of HealthConcourse providing the internal FHIR server that persists FHIR resources (aka, the HealthConcourse cache) and provides the APIs and access mechanisms exposed to internal services as well as external systems of engagement for data retrieval. HealthConcourse includes multiple instances of Smile CDR optimized for different purposes (e.g., storing and managing clinical records, storing and managing administrative and product configuration data, storing and managing message passing interface (MPI) data, etc.).

- **Perspecta knowledge manager**: Perspecta teamed up with Dynamic Content Group (DCG)—who employs recognized leaders in the emerging field of standards-based, executable knowledge artifacts and knowledge expression languages such as Clinical Query Language (CQL)–to build the knowledge manager component of HealthConcourse. The knowledge manager is a revolutionary, first of its kind, knowledge broker that dynamically discovers and evokes knowledge services to evaluate and learn from a patient’s clinical records. Knowledge services are computable expressions of knowledge (e.g., clinical decision support rules, quality metrics, machine learning models, statistic computations, medical calculators, algorithms, decision trees, etc.) delivered as microservices with standards-based APIs. The most commonly used API standard for knowledge services is CDS-Hooks. By leveraging standards, the knowledge manager works with any number of knowledge services and dynamically adds or removes knowledge services as needed. The knowledge manager maintains a registry of knowledge services and, through the knowledge service discovery process, learns and maintains information about the data requirements, pre-conditions, usage context/requirements and outputs of each knowledge service. Armed with this information, HealthConcourse determines if new data ingested and persisted into the HealthConcourse cache is appropriate and sufficient to evoke a knowledge service. The knowledge services run their algorithm against the data provided and return new information (e.g., risk assessments, computed results, guidance, recommendations, etc.) that is added as new FHIR
resources into the patient's aggregated data set. This results in "closed-loop analytics" that dramatically reduces the time it takes to
get knowledge, and the information resulting from knowledge services, into the hands of clinicians, patients and decision makers.
The knowledge manager is an important step toward the democratization of knowledge: freeing important business rules, models and
algorithms from proprietary and opaque systems and enabling a federated, standards-based knowledge community where knowledge
is authored external to record management systems (e.g., EMRs) by numerous people and organizations and ingested and leveraged
within robust data platforms like HealthConcourse.

**Red Hat® Process Automation Manager:** The Red Hat Process Automation Manager (PAM) is the run-time BPM+ engine. BPM+ is an
evolution of the popular business process modeling notation (BPMN) standard historically used for workflow and process engineering.
BPM+ extends BPMN by adding additional standards for decision modeling and case management modeling. When BPM+ is applied
to health care, clinical protocols, pathways and guidelines traditionally delivered through non-digital means, and not integrated into
the tooling used by health care professionals, it can be reimagined as computable expressions of process knowledge. By modeling
a clinical pathway (e.g., the sequence of activities and decisions recommended to treat a common condition like hypertension) as
a BPM+ model, the pathway can be implemented as software with automation of tasking and workflow. HealthConcourse takes this
one step further. By leveraging and extending the PAM, HealthConcourse uses BPM+ models to know exactly when certain data and/
or knowledge services could be used within the context of a clinical workflow. This allows us to convert the menial human tasks of
finding specific data in a patient's longitudinal health record and performing manual calculations or inferences into an automated
delivery of FHIR data and standards-based knowledge services. In essence, the HealthConcourse implementation of the PAM reduces
the cognitive burden on clinical end-users while simultaneously aligning clinical activities to the evidence and best practice expressed
in the BPM+ models.

**Red Hat® 3scale API Manager:** HealthConcourse is an API based platform. The 3scale API Manager helps to organize, document and
control access to APIs. HealthConcourse uses an API manager as a key part of the access gateway used by systems of engagement to
access patient data and HealthConcourse services. The API manager helps enforce security controls and helps monitor API usage.

**HealthConcourse services ecosystem**

HealthConcourse is based on a service-oriented architecture underpinned by the core platform. Specific, discrete functionality
is encapsulated in microservices and accessed through predictable, standards-based APIs. HealthConcourse takes a 3-layered
encapsulation and abstraction approach to integrating microservices into the platform.

- **Layer 1:** NiFi is used to define the orchestration flow for microservices and control the inputs and outputs. Each time an external
  microservice is required, a NiFi process is created to manage the interactions with the microservice and incorporate it into the
  larger orchestration flow
- **Layer 2:** HealthConcourse aims to enable a plug-and-play approach to technology integration to maximize portability, leverage
  best-of-breed, and reduce/manage vendor-lock. For each microservice required by HealthConcourse, a service wrapper is created
to encapsulate and manage the underlying technology that satisfies the requirements of the microservice. The service wrapper
API is exposed to NiFi (and other consumers of the microservice) to abstract the consumers of the microservice from the specific
technology "under the hood" for a given build at a specific point in time
- **Layer 3:** The best technology for the job is selected for each microservice. Preference is given to technologies that meet
  all functional requirements, have high performance and throughput, are stateless and proven to scale and enforce rigorous
  security controls. HealthConcourse is built on top of a pre-selected set of technologies (as shown in figure 3), but because
each technology is encapsulated behind a service wrapper and NiFi processor, the technologies can change based on specific
customer requirements or new advances in technology

The actual number of microservices needed by HealthConcourse changes over time as HealthConcourse evolves. The number of
microservices is also variable based on specific implementation requirements (not all microservices are needed for all customers).
That said, there are a common set of microservices that make up the bulk of the HealthConcourse services ecosystem.

**Transformation services:** Microservice(s) that map and transform data structures and formats provided by systems of record into FHIR
resources. The default and preferred technology used for this service is the open source MDMI product. Smile CDR and NiFi are also
used for FHIR transforms in some situations.

**Terminology services:** Microservice(s) that map and translate coded clinical concepts (i.e., clinical concepts that are bound to
standard medical terminologies like SNOMED CT) to the preferred text description and code for that concept. Preferred terminology
bindings are determined by the FHIR specification or through FHIR profiles. The default and preferred technology used for this service
is the Wolters Kluwer Health Language terminology server. The UMLS Metathesaurus and the National Library of Medicine (NLM) Value
Set Authority Center (VSAC) value sets are also used in some situations.

**Identity management services:** Microservice(s) that manage identities and resolve identity differences for entities—including
people (e.g., master person index [MPI]). Probability-based matching algorithms are commonly used to quantify the likelihood that
information from multiple sources does or does not represent the same person (or other physical entity). The default technology used
for this service is the Smile CDR MPI. Provider and organization directories also fit into this category of services.
Unstructured data/NLP services: Microservice(s) that use natural language processing to mine free-text data to identify and extract important clinical concepts such as problems, observations, tests, allergies, medications, etc. Extracted concepts are mapped to standard terminologies and turned into new FHIR resources added to the patient’s integrated data set. This effectively converts data that was previously only human readable into computable, machine readable data that is useful for analytics- and algorithmic-based computations. The default technology used for this service is the Perspecta consent management and filtering solution.

Data validation services: Microservice(s) that validate FHIR resources against the FHIR specification and against FHIR profiles to identify issues and measure the quality of the data. Using FHIR profiles for validation allows us to add, edit or remove profiles on demand to dynamically configure the validation rules. The default and preferred technology used for this service is the open source cTAKES™ library with health care specific extensions developed in collaboration with The University of Utah.

Data labeling and metadata services: Microservice(s) that add metadata to the FHIR to improve the utility and understanding of the data. The data labeling service, for instance, analyzes records against data sensitivity categories and adds metadata tags to the FHIR resource indicating if the contents are sensitive or restricted. Doing this allows us to implement fine-grained consent management whereby patients can decide exactly who can see specific portions of their longitudinal records and for what purpose. The default technology used for this service is the Wolters Kluwer Health Language Code Group Manager.

Deduplication services: Microservice(s) that identify record duplicates that invariably arise when clinical records are aggregated from the same patient across different sites. FHIR Linkage resources are created to capture information related to the duplications between records. The default technology used for this service is the deduplication and reconciliation product built in collaboration between Perspecta and Amida.

Record linking services: Microservice(s) that identify record linkages that occur when comparing parts of the patient’s longitudinal record against knowledge bases to identify associative relationships such as which drugs in the patient's medication list treat which conditions in the patient's problem list. Linkage resources are created that capture the associative information. The default technology used for this service is the Precision Medicine Summary knowledge base from Intelligent Medical Objects (IMO).

Privacy and consent services: Microservice(s) that create and manage patient consent provisions and filter FHIR query results based on consent rules. This capability leverages the sensitivity tagging from the data labeling and metadata services. The default technology used for this service is the Perspecta consent management and filtering solution.

Access control services: Microservice(s) that enforce security controls. Access control services may not strictly be implemented as microservices the way other services are. The Open Authorization (Oauth) framework is used as part of this service.

HealthConcourse infrastructure

HealthConcourse infrastructure is organized into three primary categories: cloud hosting, containerization and security.

Cloud hosting: HealthConcourse is a cloud native platform, but it is intentionally cloud-agnostic. To be cloud-agnostic, HealthConcourse is built on technologies that are self-contained and do not rely on proprietary cloud provider features. HealthConcourse leverages cloud compute, security and network capabilities to assist in scalability, access and cybersecurity. HealthConcourse avoids cloud capabilities for storage and application functionality to avoid vendor-lock. HealthConcourse has been built, tested and deployed in AWS and Microsoft Azure, the two most prominent cloud platforms on the market today.

Containerization: HealthConcourse is completely containerized using Docker and orchestrated using Docker Enterprise Kubernetes supplied by the OpenShift Container Platform (OCP) and OpenShift. This provides cloud portability and enables monitoring, security and built-in scalability.

Security: Our DevSecOps process starts with our container platform. OCP affords HealthConcourse a container platform that provides enterprise Kubernetes and enforces that all containers run as a non-privileged user (not ROOT). Furthermore, each container is backed by an enterprise operating system that provides Red Hat’s full support and attention to tracking and fixing critical security vulnerabilities (CVEs) in cooperation with the National Institute of Standards and Technology (NIST) security checklists via OpenSCAP (Compliance as Code). Container security and runtime security is further enforced using Twistlock. Twistlock provides pipeline to perimeter security via vulnerability management and compliance by scanning all containers through the development pipeline and defense in depth during runtime as it monitors access control and employs machine learning to uncover threats. Vulnerabilities are identified not only by Twistlock, but also during code compilation and deployment via static code analysis, OWASP, and other tools. HealthConcourse also leverages inherent network and infrastructure security controls provided by the cloud provider. HealthConcourse can handle personally identifiable information (PII) and personal health information (PHI) securely and be in compliance with all relative privacy and security policies.

4. HealthConcourse demo environment

For demonstration and validation purposes, we built a robust demo environment to mimic the complexity and heterogeneity of a real-world health IT computing environment. Figure 4 shows the additional solutions that make up the HealthConcourse demo environment.
The demo environment includes four different categories of knowledge services: proprietary, closed, black-box systems. The knowledge services are typically expressed as clinical decision support rules, business rules and/or analytics models historically buried inside an unlimited number of knowledge services. This is hugely important for unlocking and democratizing health care knowledge.

Knowledge services: The HealthConcourse knowledge manager can dynamically incorporate (register, discover and invoke) an unlimited number of knowledge services. This is hugely important for unlocking and democratizing health care knowledge commonly expressed as clinical decision support rules, business rules and/or analytics models historically buried inside proprietary, closed, black-box systems.

The demo environment includes four different categories of knowledge services:

1. Motive knowledge services: Motive ML, one of our knowledge engineering partners, delivers clinical decision support rules through standards-based knowledge services leveraging the CDS Hooks standard. Upon registration with HealthConcourse, the knowledge manager discovers the required data, pre-conditions, usage context and output necessary to fully integrate and dynamically invoke the motive knowledge services. Outcomes of CDS Hooks-style knowledge services are cards that get transposed into FHIR GuidanceResponse resources.

2. DCG quality measures: Quality measures are another form of knowledge services. DCG transformed industry quality measures into computable expressions in CQL. Integration occurs through the HealthConcourse knowledge manager.

3. Medal knowledge services: Medal, The Medical Algorithms Company, has amassed a collection of more than 20,000 medical calculators based on exhaustive literature review over many years. With emerging standards like CDS Hooks, these calculators can now be delivered as software able to integrate natively through our knowledge manager.

4. Microsoft machine learning (ML) models: The ML models from Microsoft are typically used for population health analytics (see description below). However, once a ML model has been trained to identify patient risk for disease based on a variety of confounding factors, the ML model itself can be delivered as-a-services accessed through a CDS Hooks interface. This allows ML models to natively integrate into HealthConcourse to evaluate patient’s risk.

Systems of record: The HealthConcourse demo environment includes data from a variety of data sources. Specific EMR systems were chosen for integration because they are representative of those used by the Veterans Health Administration (VHA) and the Defense Health Agency (DHA). The Veterans Health Information Systems and Technology Architecture (VistA) system is the VHA’s legacy EMR and Cerner Millennium has been selected as the future EMR for the VHA and DHA. Both are primary sources of data for HealthConcourse. The demo environment includes the open-source version of VistA and the Cerner Millennium sandbox. VistA and Cerner represent the legacy and future EMR, respectively, in use and/or planned for the VA. Indeed, the demo environment includes several running instances of VistA to recreate some of VA’s environmental complexity around managing multiple, parallel VistA systems. Some of these VistA systems are accessed directly through VistA remote procedure calls and other instances are accessed via the InterSystems HealthShare platform—the VA’s preferred technology for managing VistA data access going forward. In addition, an Allscripts sandbox is included as it represents the type of EMR system that active military patients or veteran patients may encounter when they receive care from private care providers. Some of the data loaded into VistA is synthetic health data generated from the Synthea synthetic data generator created by the MITRE Corporation. This gives us a rich and robust data set to work without breaching patient privacy or incurring security concerns. Logica sandboxes are also included to provide even greater diversity and data federation.

Integrating with multiple EMRs allows us to recreate an important business challenge facing health care delivery today: a patient is seen by multiple providers in different organizations (e.g., VHA, DHA and private practitioners) each using different EMR products. Each EMR is essentially a boundary that stores and controls a portions of the patient data. The patient may be known to the Defense Health Agency (DHA). The Veterans Health Information Systems and Technology Architecture (VistA) system is the VHA’s legacy EMR and Cerner Millennium has been selected as the future EMR for the VHA and DHA. Both are primary sources of data for HealthConcourse. The demo environment includes the open-source version of VistA and the Cerner Millennium sandbox. VistA and Cerner represent the legacy and future EMR, respectively, in use and/or planned for the VA. Indeed, the demo environment includes several running instances of VistA to recreate some of VA’s environmental complexity around managing multiple, parallel VistA systems. Some of these VistA systems are accessed directly through VistA remote procedure calls and other instances are accessed via the InterSystems HealthShare platform—the VA’s preferred technology for managing VistA data access going forward. In addition, an Allscripts sandbox is included as it represents the type of EMR system that active military patients or veteran patients may encounter when they receive care from private care providers. Some of the data loaded into VistA is synthetic health data generated from the Synthea synthetic data generator created by the MITRE Corporation. This gives us a rich and robust data set to work without breaching patient privacy or incurring security concerns. Logica sandboxes are also included to provide even greater diversity and data federation.

Knowledge services: The HealthConcourse knowledge manager can dynamically incorporate (register, discover and invoke) an unlimited number of knowledge services. This is hugely important for unlocking and democratizing health care knowledge commonly expressed as clinical decision support rules, business rules and/or analytics models historically buried inside proprietary, closed, black-box systems.

The demo environment includes four different categories of knowledge services:

1. Motive knowledge services: Motive ML, one of our knowledge engineering partners, delivers clinical decision support rules through standards-based knowledge services leveraging the CDS Hooks standard. Upon registration with HealthConcourse, the knowledge manager discovers the required data, pre-conditions, usage context and output necessary to fully integrate and dynamically invoke the motive knowledge services. Outcomes of CDS Hooks-style knowledge services are cards that get transposed into FHIR GuidanceResponse resources.

2. DCG quality measures: Quality measures are another form of knowledge services. DCG transformed industry quality measures into computable expressions in CQL. Integration occurs through the HealthConcourse knowledge manager.

3. Medal knowledge services: Medal, The Medical Algorithms Company, has amassed a collection of more than 20,000 medical calculators based on exhaustive literature review over many years. With emerging standards like CDS Hooks, these calculators can now be delivered as software able to integrate natively through our knowledge manager.

4. Microsoft machine learning (ML) models: The ML models from Microsoft are typically used for population health analytics (see description below). However, once a ML model has been trained to identify patient risk for disease based on a variety of confounding factors, the ML model itself can be delivered as-a-services accessed through a CDS Hooks interface. This allows ML models to natively integrate into HealthConcourse to evaluate patient’s risk.
**BPM+ models and supporting solutions:** BPM+ for health is a very exciting development in health informatics. BPM+ models allow quality organizations and process models to represent the evidence-based best practice for practicing medicine in machine-readable, automatable clinical workflows and pathways. Instead of producing hundreds of pages of documentation explaining how to treat complex medical conditions, the knowledge baked into the clinical tasking, workflow, data requirements, decision points and end conclusions can be expressed in a BPM+ model. HealthConcourse is leading the way in implementing and automating BPM+ models as running software. In essence, a BPM+ model is an expression of four related things: tasking/activities, sequence, decisions and required data (inputs and outputs into tasks, activities and decisions). When a BPM+ model is integrated into HealthConcourse, all four of these items can be automated. Automation provides two important benefits: the ability to enforce the clinical best practices and knowledge expressed in the model—without ambiguity or unnecessary or undesired variance—and the ability to reduce the cognitive burden on clinicians by automating data identification (e.g., chart review to find specific information), data delivery (contextually relevant within the appropriate activity or decision in the model), decision support (through knowledge services) and task/activity sequencing. The demo environment includes the following solution components.

1. Trisotech’s BPM+ modeling tool: Trisotech is a market leader in BPM+ tooling, particularly for creating robust, implementable models. For the demo environment, the Trisotech modeling tool was used to reproduce clinical pathways published by the VA and DOD for the treatment of chronic kidney disease and diabetes as computable, well-formed, BPM+ models. The modeling environment comes with a graphical editor with drag and drop ability to generate, link and organize activities, decisions, data objects and other relevant modeling widgets into a comprehensive workflow model. Data objects are bound as input and output parameters to modeled activities, conditions and decisions. Knowledge and decision support is also included and bound to the model either through decision modeling notation (DMN) expressions or through service calls to externally defined knowledge services. The models produced in the Trisotech tooling can be platform independent models, in that they do not bind the data and knowledge endpoints of any one particular target platform. While HealthConcourse provides FHIR endpoints for data and dynamically incorporated knowledge services for decision support, HealthConcourse may not be the only platform that runs a given model. To ensure platform portability and maximize uptake and reuse of BPM+ models, we used Trisotech to produce platform independent models to the greatest extent possible.

2. A model operationalization environment powered by MDM: The model operationalization step is necessary to take a platform independent model and add the additional metadata and platform specific bindings that allow it to be integrated and run in a target implementation environment. This includes binding the data objects to actual data retrieval queries or APIs. For HealthConcourse, this means mapping the data objects to the HealthConcourse FHIR APIs. An intermediary called the MDM Semantic Element Exchange Repository (SEER) provides platform independent data objects to the modeler in the Trisotech modeling environment and allows for platform specific binding to HealthConcourse FHIR APIs upon operationalization. By using the SEER, the data objects included in a platform independent model can be mapped to different platform specific implementations. Operationalization also includes mapping knowledge service calls expressed in the model to actual knowledge services running in HealthConcourse or other target platforms. Platform specific knowledge binding is accomplished when the BPM+ model is registered and run through the model discovery process within HealthConcourse. The model registration user interface (UI) provides the model registrant with a list of available knowledge services. When the registrant binds a decision node in the BPM+ model to an actual knowledge service, transformation logic can be added in the form of CQL scripts that transforms the model input and output parameters to the parameters used by the knowledge service. In the future, as CDS Hooks and other knowledge standards evolve, we anticipate that this manual step can be automated by unambiguously expressing the knowledge and decision support semantics required by the BPM+ model such that knowledge service binding can occur through an automated discovery and matching process.

3. The BPM+ models registered and implemented as microservices: After operationalization, BPM+ models are compiled as knowledge services and metadata for the models (including data and knowledge bindings) and are stored as FHIR resources in HealthConcourse. The BPM+ model microservices are automated when they are invoked by HealthConcourse-mediated processes or through the HealthConcourse viewer UI (see below). Automation includes the delivery of data and knowledge endpoints as well as the execution of the workflow and orchestration of modeled activities.

**Systems of engagement:** The demo environment also includes several systems of engagement sitting on top of HealthConcourse. The objective of these systems is to demonstrate how HealthConcourse supports multiple use cases implemented by different apps and platforms developed by different companies and developers. This is possible when the apps share a common set of data requirements and have the ability to access this data through standards-based APIs. The systems of engagement layered in the demo environment include several UI applications and platforms developed by Perspecta and our partners. These include:

1. A general-purpose patient viewer: The patient viewer enables users to see the aggregated, standardized and enriched data for a select patient via the FHIR APIs. The patient viewer shows the core clinical and demographic data obtained from the systems of record as well as the inferred data derived from evaluating the source data with knowledge bases added into the data processing pipeline. The patient viewer includes a BPM+ enabled workflow visualizer. The BPM+ user interface depicts a clinical workflow as an interactive process model. This embedded UI highlights current activities and navigation across activities defined by the business process. It demonstrates the power of BPM+ models and event-driven architecture as many activities that currently require manual intervention and cognitive burden are now automated.

2. MSFT machine learning (ML) dashboards: Leveraging the Microsoft Azure Machine Learning environment, Microsoft and Perspecta collaborated to build a “chronic kidney disease” and a “diabetes prediction” ML model and dashboard that uses machine learning
to stratify a population of patients into low, medium and high risk cohorts. These ML models acquire population data from HealthConcourse through our FHIR Batch API

3. PRA Clinical6: The PRA Clinical6 app is used today at the VA to enroll veterans in the VA’s Million Veteran Program (MVP). Clinical6 integrates with HealthConcourse through our FHIR APIs. Aligning to this standard allowed both apps to complete integration in a couple weeks

4. CareNexus: The CareNexus app built by Book Zurman is a commercial version of the open source Omnibus Care Plan application developed by and for SAHMSA (Substance Abuse and Mental Health Services Administration). CareNexus integrates with HealthConcourse through the FHIR APIs

Previous demo builds have included several other systems of records including medical devices and patient sources as well as several more systems of engagement including SMART on FHIR apps, mobile apps and population dashboards.

5. The HealthConcourse coalition of partners

Developing the many services and capabilities within HealthConcourse requires a coalition of partners with a variety of products aligned to the needs of the DHP reference architecture. Our partners and collaborators represent a thriving and evolving community of organizations that are collectively innovating within the Perspecta development lab. Each of these companies offers cutting edge technologies in their own area of expertise. The ecosystem is greater than the sum of its parts when technologies from these companies are brought together with cohesion and orchestrated through a standards-based platform.

Figure 5. Our primary partners

6. References
