#### 环境层（需要与苏展对接）

环境层是基础平台由领域数据、数据环境、网络环境，通用环境和主机环境构成。

领域数据是我们在网信行业多年以来积累的多种信息来源、多种数据格式的海量数据。它包括从各大互联网平台采集到的音视频等文件组成原始碎片数据；对原始碎片数据进行清洗、解析和识别之后组成的结构化的特征大数据；为机器翻译提供支持的多语种语料数据；以及多年积累的领域内外的领域知识和通用知识。

为了存储多信源数据，我们所使用到的数据环境包含多种类型的数据存储。

包括关系型数据库、列式数据库、内存数据、消息中间件、应用中间件、应用中间件、搜索引擎、分布式文件系统和分布式计算架构。

我们的产品服务支持部署在互联网、网信专网，电子政务内网和工控物联网等多种网络环境下。

通用环境包含操作系统、容器环境、微服务框架、分布式监控，自动化工具和版本管理工具。

主机环境包含私有云和公有云的各种类型的服务器，包括计算服务器、存储服务器，内存服务器和阿里云服务器等。

##### 2.2.1.3.1领域数据

领域数据包括从各大互联网平台采集到的音视频等文件组成原始碎片数据以及对原始碎片清洗后的特征大数据。原始碎片包括HTML、MP4、MP3等原始数据碎片，特征大数据包括HBase中经过算法支撑识别之后的带有特征的数据。

##### 2.2.1.3.2数据环境

数据环境使用了各种类型的数据存储，其中包括关系型数据库、列式数据库、内存数据、消息中间件、应用中间件、应用中间件、搜索引擎、分布式文件系统和分布式计算架构。包括以MySQL Galera为代表的关系型数据库集群、列式数据库HBase、内存数据库Redis、图数据库Neo4j和Janus graph、时序数据库Prometheus、消息中间件Kafka、应用中间件tomcat和nginx、搜索引擎elastic search以及分布式文件系统FastDFS和分布式计算框架Spark，实现对文本、视频、图片、音频等多种类型的非结构化、半结构化和结构化数据的存储以及生命周期的管理。

##### 2.2.1.3.3网络环境

我们的产品服务部署在互联网、网信专网，电子政务内网和工控物联网等多种网络环境下。

通过互联网收集数据，传递信息提供给客户。

##### 2.2.1.3.4通用环境

通用环境包括操作系统、容器环境、微服务框架、分布式框架、自动化工具和版本管理工具。操作系统主要使用CentOS和RancherOS；使用kubernetes实现了对容器的集群式编排和调度，使用Rancher管理kubernetes集群，部署服务项目等;通过无侵入并且语言无关的微服务框架实现对微服务的精细化管理；通过分布式监控框架实现复杂网络环境下多集群多主机多容器的性能监控和预警；通过自动化工具实现DevOps从计划、编码、构建、测试、发布、部署、运维、监控的全生命周期的规划。

##### 2.2.1.3.5主机环境

主机环境包括计算服务器、存储服务器、内存服务器和阿里云服务器。

计算服务器主要支撑Spark服务和算法服务等计算密集型服务，存储服务器主要假设MySQL、HBase、FastDFS等存储服务，阿里云服务器主要支撑组装层服务。