Seven features of big data:

1.Velocity 2. Volume 3. Variety 4. Value 5. Veraacity6. Viusualization 7. Variability

System Development Lifecycle (SDLC) element:

1.Analysis 2.Design 3.Implementation 4.Testing 5.Maintenance 6.Planning

Typical – Waterfall

**Definition:**

A typical approach of SDLC

Step by Step as started from Planning

**Advantage:**

1.Scaled process

2.Well defined requirements

3.Disadvantage

4.The development is not easy to change Once the planning is confirmed

V-model:

图示

描述已自动生成

System development Life Cycle (SDLC) and System Testing Life Cycle (STLC) doing work in parallel.

To solve the issue of Waterfall model:

testing in the model starts only after implementation is done.

|  |  |
| --- | --- |
| Advantage | Disadvantage |
| This is a highly-disciplined model and Phases are completed one at a time. | Not a good model for complex and object-oriented projects. |
| Works well for smaller projects where requirements are very well understood. | Poor model for long and ongoing projects |
| Easy to manage due to the rigidity of the model. Each phase has specific deliverables and a review process. | Not suitable for the projects where requirements are at a moderate to high risk of changing. |

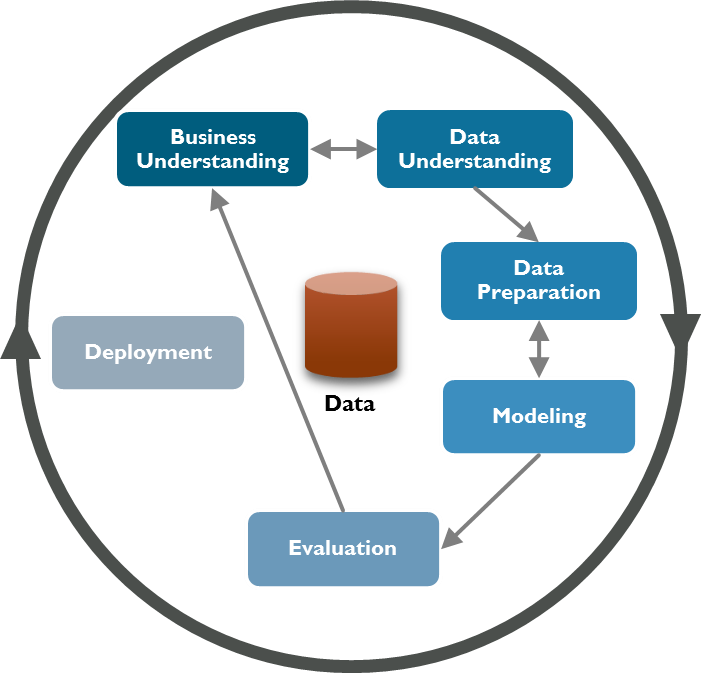
Comparison Between WaterFall and V-model

|  |  |  |
| --- | --- | --- |
| Category | WaterFall | V-model |
| Development Stages | Planning, Analysis, Design, Coding/Implementation, Testing, Maintenance | Verification: Business analysis, System analysis, High Level, Low Level, Coding.  Validation: Unit, Integration, system, appearance testing |
| Flow | Once you go, Never Come back. | Once you go, Never Come back until you do validation |
| When | A large scale project  A clear requirement and objective | A large scale project  A clear requirement and objective |
| Test | Just a part of process | a series of tests for different levels of development |

Elements in Lifecycle of Big Data Project

|  |  |  |
| --- | --- | --- |
| Business Case Evaluation | Data Identification | Data Acquisition & Filtering |
| Data Extraction | Data Validation & Cleansing | Data Aggregation & Representation |
| Data Analysis | Data Visualization | Utilization of Analysis Results |

data mining (Crisp-DM)



But There are some limitations:

Some stages cannot backward

Difficult to track the progress of project

The division of work is clear but cannot support each other

Otherwise, the whole team have to do the same thing as the same stage.

Version of Data become messy and complicated

Agile focus on the following concerns:

1.Individuals and communications

2.Co-operate with stakeholder

3.Response to changes

4.Working software over comprehensive documentation

Agile development principles (Manifesto)

* Customer satisfaction by early and continuous delivery of valuable software.
* Welcome changing requirements, even in late development.
* Deliver working software frequently (weeks rather than months)
* Close, daily cooperation between business people and developers
* Projects are built around motivated individuals, who should be trusted
* Face-to-face conversation is the best form of communication (co-location)
* Working software is the primary measure of progress
* Sustainable development, able to maintain a constant pace
* Continuous attention to technical excellence and good design
* Simplicity—the art of maximizing the amount of work not done—is essential
* Best architectures, requirements, and designs emerge from self-organizing teams
* Regularly, the team reflects on how to become more effective, and adjusts accordingly

Scrum

Three position: Development Team,Product Owner,Scrum Master

表格

描述已自动生成