ODI Interview Questions

1. [**Explain what is ODI? Why is it different from the other ETL tools?**](https://mindmajix.com/odi-interview-questions#odi)
2. [**What is E-LT?**](https://mindmajix.com/odi-interview-questions#elt)
3. [**What is Oracle Data Integration Suite?**](https://mindmajix.com/odi-interview-questions#data-integration-suite)
4. [**Does ODI support web services?**](https://mindmajix.com/odi-interview-questions#web-services)
5. [**What is a profile in ODI?**](https://mindmajix.com/odi-interview-questions#profile)
6. [**How will you bulk load data?**](https://mindmajix.com/odi-interview-questions#bulk-load-data)
7. [**What is a Procedure?**](https://mindmajix.com/odi-interview-questions#procedure)
8. [**What is a Scenario?**](https://mindmajix.com/odi-interview-questions#scenario)
9. [**What are the types of Knowledge Modules?**](https://mindmajix.com/odi-interview-questions#knowledge-modules)
10. [**How to write the sub-queries in ODI?**](https://mindmajix.com/odi-interview-questions#sub-queries)
11. [**What are the prime responsibilities of the Data Integration Administrator?**](https://mindmajix.com/odi-interview-questions#data-integration-administrator)

## Basic ODI Interview Questions and Answers

### **1. What is Oracle Data Integrator (ODI)?**

Oracle acquired Sunopsis in 2006 and with it “Sunopsis Data Integrator”.

[**Oracle Data Integrator (ODI)**](https://mindmajix.com/what-is-odi) is an E-LT (Extract, Load and Transform) tool used for high-speed data movement between disparate systems.

The latest version, Oracle Data Integrator Enterprise Edition (ODI-EE) brings together “Oracle Data Integrator” and “Oracle Warehouse Builder” as separate components of a single product with a single license.

### **2. What is E-LT?**

E-LT is an innovative approach to extracting, loading, and Transforming data. Typically ETL application vendors have relied on costly heavyweight, mid-tier servers to perform the transformations required when moving large volumes of data around the enterprise.

ODI delivers unique next-generation, Extract Load and Transform (ELT) technology that improves performance and reduces data integration costs, even across heterogeneous systems by pushing the processing required down to the typically large and powerful database servers already in place within the enterprise.

### **3. What components makeup Oracle Data Integrator?**

“Oracle Data Integrator” comprises of:

1) Oracle Data Integrator + Topology Manager + Designer + Operator + Agent

2) Oracle Data Quality for Data Integrator

3) Oracle Data Profiling

### **4. What is Oracle Data Integration Suite?**

Oracle data integration suite is a set of data management applications for building, deploying, and managing enterprise data integration solutions:

1. Oracle Data Integrator Enterprise Edition
2. Oracle Data Relationship Management
3. Oracle Service Bus (limited use)
4. Oracle BPEL (limited use)
5. Oracle WebLogic Server (limited use)

Additional product options are:

1. Oracle Goldengate
2. Oracle Data Quality for Oracle Data Integrator (Trillium-based DQ)
3. Oracle Data Profiling (Trillium based Data Profiling)
4. ODSI (the former Aqualogic Data Services Platform)

### **5. What systems can ODI extract and load data into?**

ODI brings true heterogeneous connectivity out-of-the-box, it can connect natively to Oracle, Sybase, MS SQL Server, MySQL, LDAP, DB2, PostgreSQL, Netezza.

It can also connect to any data source supporting JDBC, its possible even to use the Oracle BI Server as a data source using the JDBC driver that ships with BI Publisher

### **6. What are Knowledge Modules?**

 Knowledge Modules form the basis of ‘plug-ins’ that allow ODI to generate the relevant execution code, across technologies, to perform tasks in one of six areas, the six types of knowledge module consist of:

* Reverse-engineering knowledge modules are used for reading the table and other object metadata from source databases.
* Journalizing knowledge modules record the new and changed data within either a single table or view or a consistent set of tables or views.
* Loading knowledge modules are used for efficient extraction of data from source databases for loading into a staging area (database-specific bulk unload utilities can be used where available).
* Check knowledge modules are used for detecting errors in source data.
* Integration knowledge modules are used for efficiently transforming data from the staging area to the target tables, generating the optimized native SQL for the given database.
* Service knowledge modules provide the ability to expose data as Web services.
* ODI ships with many knowledge modules out of the box, these are also extendable, they can be modified within the ODI Designer module.

### **7. How do ‘Contexts’ work in ODI?**

ODI offers a unique design approach through the use of Contexts and Logical schemas. Imagine a development team, within the ODI Topology manager a senior developer can define the system architecture, connections, databases, data servers (tables, etc), and so forth.

These objects are linked through contexts to ‘logical’ architecture objects that are then used by other developers to simply create interfaces using these logical objects, at run-time, on the specification of a context within which to execute the interfaces, ODI will use the correct physical connections, databases + tables (source + target) linked the logical objects being used in those interfaces as defined within the environment Topology.

### **8. Does my ODI infrastructure require an Oracle database?**

No, the ODI modular repositories (Master + and one of the multiple Work repositories) can be installed on any database engine that supports ANSI ISO 89 syntax such as Oracle, Microsoft SQL Server, Sybase AS Enterprise, IBM DB2 UDB, IBM DB2/40.

### **9. Where can I get more information on ODI?**

The OTN Data integration home page:**https://bit.ly/1Myxa4J**

### **10. Does ODI support web services?**

Yes, ODI is ‘SOA’ enabled and its web services can be used in 3 ways:

* The Oracle Data Integrator Public Web Service, lets you execute a scenario (a published package) from a web service call.
* Data Services, which provide a web service over an ODI data store (i.e. a table, view, or other data source registered in ODI).
* The ODIInvokeWebService tool that you can add to a package to request a response from a web service.

### **11. Where does ODI sit with my existing OWB implementation(s)?**

 As mentioned previously, the ODI-EE license includes both ODI and OWB as separate products, both tools will converge in time into “Oracle’s Unified Data Integration Product”.

Oracle has released a statement of direction for both products, published January 2010:

OWB 11G R2 is the first step from Oracle to bring these two applications together, it's now possible to use ODI Knowledge modules within your OWB 11G R2 environment as ‘Code

Templates’, an Oracle white paper published in February 2010 describes this in more detail.

### **12. What is the ODI Console?**

ODI console is a web-based navigator to access the Designer, Operator, and Topology components through the browser.

### **13. Suppose I having 6 interfaces and running the interface 3 rd one failed how to run the remaining interfaces?**

If you are running Sequential load it will stop the other interfaces. so go to the operator and right-click on the filed interface and click on restart.

If you are running all the interfaces are parallel only one interface will fail and other interfaces will finish.

### **14. What are load plans and types of load plans?**

A load plan is a process to run or execute multiple scenarios as a Sequential or parallel or conditional based execution of your scenarios.

And same we can call three types of load plans, Sequential, parallel, and Condition-based load plans.

### **15. What is a profile in ODI?**

The profile is a set of objective-wise privileges. We can assign these profiles to the users. Users will get the privileges from the profile

### **16. How to write the sub-queries in ODI?**

Using the Yellow interface and sub-queries option we can create subqueries in ODI. or Using  VIEW we can go for subqueries Or Using ODI Procedure we can call direct database queries in ODI.

### **17. How to remove the duplicate in ODI?**

Use DISTINCT at the IKM level. it will remove the duplicate rows while loading into the target.

### **18. Suppose having unique and duplicates but I want to load a unique record one table and duplicates one table?**

Create two interfaces or once the procedure and use two queries one for Unique values and one for duplicate values.

### **19. How to implement data validations?**

Use Filters & Mapping Area AND Data Quality related to constraints use CKM Flow control.

### **20. How to handle exceptions?**

Exceptions In packages advanced tab and load plan exception tab we can handle exceptions.

### **21. In the package, one interface got failed how to know which interface got failed if we no access to an operator?**

Make it mail alert or check into SNP\_SESS\_LOg tables for session log details.

### **22. How to implement the logic in procedures if the source side data deleted that will reflect the target side table?**

User this query on Command on target Delete from Target\_table where not exists (Select ‘X’ From Source\_table Where Source\_table.ID=Target\_table.ID).

### **23. If the Source has a total of 15 records with 2 records are updated and 3 records are newly inserted at the target side we have to load the newly changed and inserted records?**

Use IKM Incremental Update Knowledge Module for Both Insert n Update operations.

### **24. Can we implement a package in the package?**

Yes, we can call one package into other packages.

### **25. How to load the data with one flat file and one RDBMS table using joins?**

Drag and drop both File and table into source area and join as in Staging area.

### **26. If the source and target are oracle technology tell me the process to achieve this requirement(interfaces, KMS, Models)?**

Use LKM-SQL to SQL or LKM-SQL to Oracle, IKM Oracle Incremental update or Control append.

### **27. What we specify the in XML data server and parameters for connecting to the XML file?**

Filename with location: F and Schema :S this two parameters

### **28. How to reverse engineer views(how to load the data from views)?**

In Models Go to the Reverse engineering tab and select the Reverse engineering object as VIEW.

### **29. Is ODI Used by Oracle in their products?**

Yes there are many Oracle products that utilize ODI, but here are just a few:

1. Oracle Application Integration Architecture (AIA)

2. Oracle Agile products

3. Oracle Hyperion Financial Management

4. Oracle Hyperion Planning

5. Oracle Fusion Governance, Risk & Compliance

6. Oracle Business Activity Monitoring

Oracle BI Applications also uses ODI as its core ETL tool in place of Informatica, but only for one release of OBIA and when using a certain source system.

### **30. ELT Vs ETL**

The ability to dynamically manage a staging area:

The ability to generate code on source and target systems alike, in the same transformation.

The ability to generate native SQL for any database on the market - most ETL tools will generate code for their own engines, and then translate that code for the databases - hence limiting their generation capacities to their ability to convert proprietary concepts.

The ability to generate DML and DDL, and to orchestrate sequences of operations on the heterogeneous systems.

### **31. Explain what is ODI? why is it different from the other ETL tools?**

ODI stands for Oracle Data Integrator. It is different from another ETL tool in the way that it uses the E-LT approach as opposed to the ETL approach. This approach eliminates the need for the exclusive Transformation Server between the Source and Target Data server.

The power of the target data server can be used to transform the data. i.e. The target data server acts as a staging area in addition to its role in the target database.

While loading the data in the target database (from the staging area) the transformation logic is implemented.

Also, the use of appropriate CKM (Check Knowledge Module) can be made while doing this to implement data quality requirements.

### **32. How will you bring in the different source data into ODI?**

you will have to create data servers in the topology manager for the different sources that you want.

### **33. How will you bulk load data?**

In Odi there is an IKM that is designed for bulk loading of data.

### **34. How will you bring in files from remote locations?**

We will invoke the Service knowledge module in ODI, this will help us to accesses data through a web service.

### **35. How will you handle data quality in ODI?**

There are two ways of handling data quality in Odi….the first method deals with handling the incorrect data using the CKM…the second method uses the Oracle data quality tool(this is for advanced quality options).

### **36. What are the load plans and types of load plans?**

A load plan is a process to run or execute multiple scenarios as a Sequential or parallel or conditional based execution of your scenarios. And same we can call three types of load plans, Sequential, parallel, and Condition-based load plans.

### **37. What is a profile in ODI?**

The profile is a set of objective-wise privileges. we can assign these profiles to the users. Users will get privileges from the profile.

### **38. What is the ODI console?**

ODI console is a web-based navigator to access the Designer, Operator, and Topology components through the browser.

### **39. How to write the subqueries in ODI?**

Using the Yellow interface and subqueries option we can create subqueries in ODI.

Using  VIEW we can go for subqueries Or Using ODI Procedure we can call direct DB queries in ODI.

### **40. Suppose I having 6 interfaces and running the interface 3 rd one failed how to run the remaining interfaces?**

If you are running Sequential load it will stop the other interfaces. so go to the operator and right-click on the filed interface and click on restart. If you are running all the interfaces are parallel only one interface will fail and other interfaces will finish.

### **41. How to remove the duplicate in ODI?**

Use DISTINCT at the IKM level. it will remove the duplicate rows while loading into the target.

### **42. Suppose having unique and duplicates but I want to load a unique record one table and duplicates one table?**

Create two interfaces or once the procedure and use two queries one for Unique values and one for duplicate values.

### **43. How to write the procedures in ODI?**

The procedure is a step-by-step technology code operation. you can refer What are the types of Variables?

* Global
* Project

A variable is an object that stores a single value. This value can be a string, a number, or a date. The value is stored in Oracle Data Integrator and can be updated at run-time.

The value of a variable can be updated from the result of a query executed on a logical schema. For example, it can retrieve the current date and time from a database. A variable can be created as a global variable or in a project.

Global variables can be used in all projects, while project variables can only be used within the project in which they are defined.

### **44. Where we can use variables?**

Variables can be used in all Oracle Data Integrator expressions:ü  Mapping,ü  Filters,ü  Joins,ü  Constraints,

### **45. What is a Work Repository?**

Each work repository is attached to a master repository, therefore, information about the physical connection to a work repository is stored in the master repository it is attached to.

Defining a connection to a work repository consists of defining a connection to a master repository, then selecting one of the work repositories attached to this master repository.

### **46. What is the Master Repository?**

The Master Repository is a data structure containing information on the topology of a company’s IT resources, on security, and on version management of projects and data models.

This repository is stored on a relational database accessible in client/server mode from the different modules. Generally, only one master repository is necessary.

However, in exceptional circumstances, it may be necessary to create several master repositories in one of the following cases:

Project construction over several sites not linked by a high-speed network (off-site development, for example).

The necessity to clearly separate the interfaces’ operating environments (development, test, production), including on the database containing the master repository. This may be the case if these environments are on several sites.

### **47. What is a Procedure?**

A Procedure is a reusable component that allows you to group actions that do not fit in the Interface framework. (That is load a target datastore from one or more sources). A Procedure is a sequence of commands launched on logical schemas.

It has a group of associated options. These options parameterize whether or not the command should be executed as well as the code of the commands.

### **48. What is the Model?**

An Oracle Model is a set of datastores corresponding to views and tables contained in an Oracle Schema. A model is always based on a Logical Schema. In a given Context, the Logical Schema corresponds to a Physical Schema.

The Data Schema of this Physical Schema contains the Oracle model’s tables and views.

### **49. What is a Package?**

The package is the biggest execution unit in Oracle Data Integrator. A package is made of a sequence of steps organized in an execution diagram.

### **50. What are the User Parameters?**

Oracle Data Integrator saves user parameters such as default directories, windows positions, etc. User parameters are saved in the userpref.xml file in /bin.

### **51. What is a Project?**

A project is a group of objects developed using Oracle Data Integrator.

### **52. What is Folder?**

Certain objects in a project are organized into folders and sub-folders.

### **53. What is an Interface?**

An interface consists of a set of rules that define the loading of a Datastore or a temporary target structure from one or more source Datastores.

### **54. What is Sequence?**

A sequence is a variable automatically incremented when used. Between two uses the value is persistent. The sequences are usable like variables in interfaces, procedures, steps. A sequence can also be defined outside a project (global scope), in order to be used in all projects.

### **55. What are the User Functions?**

User functions enable you to define customized functions or “functions aliases”, for which you will define technology-dependant implementations. They are usable in the interfaces and procedures.

### **56. What is the Marker?**

Elements of a project may be flagged in order to reflect the methodology or organization of the developments. Flags are defined using markers.

These markers are organized into groups and can be applied to most objects in a project.

### **57. What is a Scenario?**

When a package, interface, procedure, or variable component is finished, it is compiled in a scenario. A scenario is the execution unit for production, that can be scheduled.

### **58. What is Context?**

A context is a set of resources allowing the operation or simulation of one or more data processing applications. Contexts allow the same jobs (Reverse, Data Quality Control, Package, etc) to be executed on different databases and/or schemas.

In Oracle Data Integrator, a context allows logical objects (logical agents, logical schemas) to be linked with physical objects (physical agents, physical schemas).

### **59. What are Memos?**

A memo is an unlimited amount of text attached to virtually any object, visible on its Memo tab. When an object has a memo attached, the icon appears next to it.

### **60. What are the Sequences?**

A sequence is a variable that increments itself each time it is used. Between two uses, the value can be stored in the repository or managed within an external RDBMS table.

Oracle Data Integrator supports two types of sequences:

**Standard sequences -** whose last value is stored in the Repository.

**Specific sequences -**whose last value is stored in an RDBMS table cell. Oracle Data Integrator undertakes to read the value, to lock the row (for concurrent updates), and to update the row after the last increment.

### **61. What is a Session?**

A session is an execution (of a scenario, an interface, a package, or a procedure, …) undertaken by an execution agent. A session is made up of steps that are made up of tasks.

### **62. What are Session Tasks?**

The task is the smallest execution unit. It corresponds to a procedure command in a KM, a procedure, assignment of a variable, etc

## Advanced ODI Interview Question and Answers

### **63. Can I create more than one Master Repository in ODI?**

Yes. In general, you need only one master repository. However, it may be necessary to create several master repositories if the Project construction over several sites not linked by a high-speed network (off-site development, for example).

The necessity to clearly separate the interfaces operating environments (development, test, production), including on the database containing the master repository. This may be the case if these environments are on several sites.

### **64. What are the types of Knowledge Modules?**

* LKM(used to extract data from heterogeneous source systems (files, middleware, databases, etc.) to a staging area).
* IKM(used to integrate (load) data from staging to target tables)
* RKM(used to perform a customized reverse-engineering of data models for a specific technology. It extracts metadata from a metadata provider to ODI repository. These are used in data models.)
* JKM(used to create a journal of data modifications (insert, update and delete) of the source databases to keep track of changes. These are used in data models and used for Changed Data Capture.)
* CKM( used to check data consistency i.e. constraints on the sources and targets are not violated. These are used in the data model’s static checks and interface flow checks. Static check refers to constraints or rules defined in the data model to verify the integrity of source or application data.
* Flow check refers to declarative rules defined in interfaces to verify an application’s incoming data before loading into target tables.)

### **65. What Is an Interface?**

An interface is an object in ODI that will map the sources to target datamarts.

### **66. What is a temporary Interface (Yellow Interface)?**

The advantage of using a yellow interface is to avoid the creation of Models each time we need to use it in an interface. Since they are temporary, they are not a part of the data model and hence don’t need to be in the Model.

### **67. Explain some differences between ODI 10g and ODI 11g?**

ODI 11g provides a Java API to manipulate both the design-time and run-time artifacts of the product. This API allows you, for example, to create or modify interfaces programmatically, create your topology, perform import or export operations, launch or monitor sessions.

This API can be used in any Java SE and Java EE applications, or in the context of Java-based scripting languages like Groovy or Jython.

External Password Storage, to have source/target data servers (and contexts) passwords stored in an enterprise credential store.

External Authentication, to have user/password information stored in an enterprise identity store (e.g.: LDAP, Oracle Directory, Active Directory), and ODI authenticating against this store. These two features let you optionally store critical information in dedicated storage and not within the ODI repository.

The ODI Console may also use Oracle’s single-sign-on systems with ODI.

### **68. What is CKM and when we will use this CKM?**

Check control module is used when we are creating constraints on the target datastore. We can say that CKM is used in Data Quality control.

### **69. What is SKM and when we will use this SKM?**

SKM (Service Knowledge Module) is used to generate code required for data services. These are used in data models.

Data Services are specialized web services that enable access to application data in datastores and to the changes captured for these datastores using Changed Data Capture.

### **70. How to load data from file to file and what are the KM’s required for this requirement?**

IKM File to File

### **71. What are the types of data quality control?**

There are two ways to data quality control:

* Static: We will run the constraints on existing target data. This is done after loading the data into the target.
* Flow: We will run the constraints on incoming data. This is done before loading the data into the target.

### **72. What is a constraint?**

It is a condition which you want to apply while transferring the data from source to target.

### **73. What is the E$ table in ODI?**

Temporary Error table created by ODI. This is created by CKM.

### **74. What is the I$ table in ODI?**

This is a flow table created by IKM while integrating data in the datamart. This is a temporary table used by ODI.

### **75. What is the J$ table in ODI?**

This is where all changes are recorded. Journals contain references to the changed records along with the type of change (insert, update or delete).

### **76. What is Journalization and why we are using it in ODI?**

It is the way to implement change data capture in ODI. We use JKM for this purpose.

### **77. Explain step-by-step procedure to enable Journalization?**

The first step is to import proper JKM. After creating a model and reverse engineering we have to add the model to CDC and then we need to subscribe to the table we want. This will enable Journalization.

### **78. Does ODI support web services?**

Yes. ODI supports web services, ODI is ‘SOA’ enabled and its web services can be used in 3 ways:

The Oracle Data Integrator Public Web Service, lets you execute a scenario (a published package) from a web service call Data Services, which provides a web service over an ODI data store (i.e. a table, view or other data source registered in ODI).

The ODI Invoke Web Service tool that you can add to a package to request a response from a web service.

### **79. What is the ODI Console?**

ODI console is a web-based navigator to access the Designer, Operator, and Topology navigators through the browser.

### **80. Suppose I having 10 interfaces and running the interface 5th one failed how to run the remaining interfaces?**

If you are running Sequential load it will stop the other interfaces. so go to operator navigator and right-click on the failed interface and click on restart.

If you are running all the interfaces are parallel only one interface will fail and other interfaces will finish.

### **81. What are the load plans and types of load plans?**

A load plan is a process to run or execute multiple scenarios as a Sequential or parallel or conditional based execution of your scenarios.

And same we can call three types of load plans, Sequential, parallel, and Condition-based load plans.

### **82. How to write the sub-queries in ODI?**

We can follow anyone of the following to create a subquery.

1. Using the Yellow interface and sub-queries option we can create subqueries in ODI.

2. Using a VIEW we can go for subqueries.

3. Using ODI Procedure we can call direct database queries in ODI.

### **83. Remove the duplicate in ODI?**

Use DISTINCT, at the IKM level. it will remove the duplicate rows while loading into the target.

### **84. Suppose having unique and duplicates but I want to load a unique record one table and duplicates one table?**

Create two interfaces or once the procedure and use two queries one for Unique values and one for duplicate values.

### **85. How to implement data validations?**

Use Filters & Mapping Area AND Data Quality related to constraints use CKM Flow control.

### **86. How to handle exceptions?**

Exceptions In packages advanced tab and load plan exception tab we can handle exceptions.

### **87. In the package, one interface got failed how to know which interface got failed if we no access to an operator?**

Make it mail alert or check into SNP\_SESS\_LOG tables for session log details.

### **88. How to implement the logic in procedures if the source side data deleted that will reflect the target side table?**

User this query on Command on target Delete from Target\_table where not exists (Select ‘X’ From Source\_table Where Source\_table.ID=Target\_table.ID).

### **89. If the source has a total of 15 records with 2 records are updated and 3 records are newly inserted. Which knowledge module we should use to get these changes on the target side.**

We have to load the newly changed and inserted records Use IKM Incremental Update Knowledge Module for Both Insert n Update operations.

### **90. Can we implement a package in the package?**

Yes. we can call one package into other packages.

### **91. How to load the data with one flat file and one RDBMS table using joins?**

Drag and drop both File and table into source area and join as in Staging area.

### **92. How to reverse engineer views(how to load the data from views)?**

In Models Go to the Reverse engineering tab and select the Reverse engineering object as VIEW.

### **93. What systems can ODI extract and load data into?**

ODI brings true heterogeneous connectivity out-of-the-box, it can connect natively to Oracle, Sybase, MS SQL Server, MySQL, LDAP, DB2, PostgreSQL, Netezza.

It can also connect to any data source supporting JDBC, its possible even to use the Oracle BI Server as a data source using the JDBC driver that ships with BI Publisher

### **94. Suppose having unique and duplicates but I want to load a unique record one table and duplicates one table?**

Create two interfaces or once the procedure and use two queries one for Unique values and one for duplicate values.

### **95. What is a procedure and how to write the procedures in ODI?**

A Procedure is a reusable component that allows you to group actions that do not fit in the Interface framework (That is load a target datastore from one or more sources).

A Procedure is a sequence of commands launched on logical schemas. It has a group of associated options.

These options parameterize whether or not the command should be executed as well as the code of the commands.

### **96. What are the prime responsibilities of the Data Integration Administrator?**

1. Scheduling and executing batch jobs.

2. Configuring, starting, and stopping the real-time services

3. Adapters configuration and managing them.

4. Repository usage, Job Server configuration.

5. Access Server configuration.

6. Batch job publishing.

7. Real-time services publishing through web services.