

---

## 个人信息

姓名：么敬国

性别：男

手机：139-1033-4380

Email：yaojingguo@gmail.com

## 教育

北京理工大学，硕士，计算机应用技术，2004年4月

北京理工大学，学士，计算机科学技术，2001年7月

## 自我评价

- 10年企业级平台架构和研发经验。
- 有以下企业级系统的架构经验：企业内容管理平台，企业流程管理平台，企业主数据管理平台，企业统一用户管理平台，CRM平台，数据仓库，大数据分析平台，统一消息平台。
- 精通主流大数据平台的架构与系统建设，熟练掌握数据分析技术和数据挖掘技术。丰富的Hadoop平台（HBase, Zookeeper, Cassandra, Pig和Hive）上数据挖掘应用开发经验。
- 精于计算机系统构建理论和技术，并发系统设计和分布式系统构建专家。
- 精通Java（J2EE），C/C++，Python，Go, Lisp和Bash。熟练使用JavaScript，Haskell和Ruby。
- 丰富的主流数据库系统使用和调优经验：Oracle, DB2, SQL Server和MySQL。
- 精通Linux环境下的系统开发，精通TCP/IP协议，熟练掌握基本算法和数据结构。熟悉云计算虚拟化相关技术。

## 英语

CET6优秀，口语流利，多年和海外团队合作经验。

## 教育，认证和培训

- MIT-6.006 Introduction to Algorithms
- MIT-6.033 Computer Systems Engineering
- MIT-6.830 Database Systems
- MIT-6.828 Operating System Engineering
- MIT-6.824 Distributed Systems
- IBM Certified Solution Designer DB2 Content Manager V8

---

## 工作经历

### 海南天涯在线网络科技有限公司 (2010/08 -- 现在)

天涯大数据部

总监

Hadoop, HBase, Pig, Mahout, C++, Python

09/2012 -- 至今

在天涯企业数据仓库的基础上开发天涯大数据分析平台：

- 天涯分布式锁系统：实现部分Google Chubby分布式锁功能，实现了Paxos算法和Replicated State Machine架构。
- 开发天涯用户标签系统：根据用户在天涯上的行为，给用户打上标签，对客户进行细分。
- 开发天涯推荐引擎：对用户行为数据和内容数据进行挖掘，提供个性化的推荐服务。
- 开发天涯词库系统：业务人员使用的领域内关键词管理平台。
- 数据挖掘，信息过滤，搜索，BI和会员团队管理。

### 天涯Hadoop数据分析平台

技术经理

Java, Lisp, Hadoop, Hive, JavaScript, MySQL, Oracle ODS, Oracle ODI, Oracle BIEE

02/2011 -- 08/2012

- 架构并带领团队建设新一代基于Hadoop/Hive的企业数据仓库来替换已有基于Oracle Data Warehouse的解决方案。
- 用Hadoop实现协同过滤推荐算法。
- 调查研究Hadoop平台技术，并向Apache提交Patch。

### 天涯虚拟化平台

资深系统架构师

Xen, KVM, QEMU

12/2010 -- 01/2011

天涯虚拟化平台的前期调研和设计工作。

### Memlink Project

Committer

C, Libevent

09/2010 -- 11/2010

---

<http://code.google.com/p/memlink/>是一个列表数据存储引擎。我设计实现了主从同步部分。

## **IBM中国软件开发中心 (2004/11 -- 2010/08)**

### **ECM/BPM CoE**

技术经理

IBM ECM, IBM BPM, SPSS Data Modler, Oracle Database, LDAP

06/2010 -- 08/2010

ECM/BPM COE的职责是促进IBM ECM/BPM软件产品在中国的成长。我们的主要任务是为ECM/BPM建立重要的样板客户。我作为COE Leader参加了银行业, 保险, 制造和核电行业的架构设计方案与实施。其中保险分析项目使用SPSS进行交叉销售, 使用了关联规则和决策树。

### **大数据查询语言Jaql**

开发组长

07/2009 -- 05/2010

Java, C++, Hadoop, Pig, SPSS Statistics, R, IBM BigData

Jaql (<http://code.google.com/p/jaql>)是IBM Hadoop大数据平台上的查询语言, 实现类似Hadoop Pig的功能。在这个项目中, 我和IBM Almaden Research Center的科研人员合作。我是中国方面的technical leader, 同时是Jaql项目的committer。我设计实现了Jaql catalog server, SPSS与Jaql的集成和JSON与CSV的转换。Catalog Server提供一个key-value存储机制。SPSS与Jaql的集成集成的目标是结合SPSS的统计数据挖掘能力和Hadoop 平台的扩展性和并行计算机制。通过这个集成, 可以平行的运行某些SPSS的算法。我还实现了JSON和CSV的转换, 并做了系统性能优化, 达到了和Pig相当的性能。我还参与了VISA信用卡反欺诈项目。

### **泰康人寿主数据项目**

架构师

04/2009 -- 06/2009

IBM MDM, IBM DB2, Informatica

我参与了主数据系统的架构设计和咨询。这个项目把泰康分散的客户数据和产品数据整合在了一起。整个项目分成3个阶段。第一阶段是ETL数据加载, 第二阶段是业务系统和主数据同步, 第三阶段主数据系统开始提供对外服务。在泰康主数据平台基础之上, 泰康开发实施了泰康分析型CRM系统, 对客户进行分类和个性化营销。

### **泰康人寿 workflow 项目**

架构师/技术组长

01/2009 -- 04/2009

J2EE, IBM ECM, IBM BPM, WebSphere, IBM MQ

---

我带领了泰康人寿保全 workflow 系统的开发。开发团队有 IBM 实验室，IBM 服务部门，IBM 合作伙伴和泰康 IT 人员组成。我架构了泰康核心业务系统的企业服务抽象。我带领团队按时高标准的完成了泰康个人保险保全项目，并和泰康建立了深厚的客户关系。在项目中，我设计实现了 IBM Content Manager 和 FileNet BPM 的集成。这个集成被当作最主要的功能点放入了 IBM CM 8.4.1 发行版，并在美国 IBM IOD 大会上首发，获得非常大的反响。

### **泰康人寿影像中心项目**

架构师

08/2008 -- 12/2008

J2EE, IBM ECM, IBM MQ

参与架构泰康人寿全集团企业级影像集中管理平台，基于企业级中间件 IBM ECM 和 IBM MQ 构建该解决方案，解决了网络环境不稳定的可靠性问题。

### **Rapid Industry Solution Enabler**

开发组长

06/2006 -- 07/2008

IBM ECM, UML, WebSphere Process Server, BPEL, Eclipse Plugin 开发

Rapid Industry Solution Enabler 是一个代码生成工具，其目标是加速 IBM ECM 和 BPM 平台上的解决方案开发。这个工具从 UML 数据模型和流程模型生成数据库 schema, 流程定义, Java 代码和 UI 代码。所有这些构成了一个可以运行的 Web 应用。这是一个 IBM CDL 和 IBM Watson Research Center 的合作项目。

### **ECM (企业内容管理)**

软件工程师/开发组长

11/2004 – 05/2006

Java, J2EE, IBM Rational Functional Tester

这是一个 IBM 中国实验室和 IBM Silicon Valley Lab 的合作项目，为 IBM ECM 产品做开发和测试。我最初用 Rational Robot 和 Rational Functional Tester 做产品界面测试，用 ANT 和 JUnit 做非界面测试。我后来加入了 Records Manager Enabler (一个 ECM 组件) 的开发团队，设计实现了 FIPS 功能点。

### **北京理工大学 (09/2001 – 04/2004)**

研究方向包括 Semantic Web 和知识推理。我参加了几个基于 J2EE 的项目。

## **文章**

- Effective Java GUI automation on multiple platforms. IBM developerWorks.  
[http://www-128.ibm.com/developerworks/rational/library/05/1004\\_yao/?ca=dgr-lnxw06JavaGUI4Win-Linux](http://www-128.ibm.com/developerworks/rational/library/05/1004_yao/?ca=dgr-lnxw06JavaGUI4Win-Linux).

- 
- Integrate FileNet BPM with IBM Content Manager, Part 3: Implement the Component Integrator-based work performers. IBM developerworks. <http://www.ibm.com/developerworks/data/library/techarticle/dm-0805yao/index.html>.
  - 一个语义Web应用研究：旅游信息系统[J]。北京理工大学学报(Ei核心刊物)。
  - 语义Web系统及其实现[J]。北京理工大学学报(Ei核心刊物)。2004，24(2):145-149。
  - 基于奇异值分解的中文Ontology自动学习技术[J]。计算机工程 (Ei Page One刊物)，2003，29(9):137-139。
  - CRL：对语义Web上的Ontology表示语言DAML-OIL的一种扩充方案[J]。计算机工程与应用(核心期刊)，2003-23。
  - 一个语义Web架构及其实现[J]。计算机工程与应用(核心期)，2003-15。