

```
{  
  "project": {  
    "name": "Interconnected AI Architecture: Parameter Scaling Constructor",  
    "version": "2.0",  
    "status": "Production Ready",  
    "date": "2026-02-12",  
    "lang": "Русский/English bilingual"  
  },  
  "documents": {  
    "total_count": 6,  
    "total_pages": "242+ страниц",  
    "files": [  
      {  
        "name": "interconnected_ai_architecture_v1.md",  
        "type": "Спецификация",  
        "pages": 50,  
        "topics": [  
          "Концепция мега-дерева",  
          "Base Encoder (замороженный)",  
          "Sparsely-Gated MoE",  
          "LoRA Adapters",  
          "Параметровые траектории масштабирования"  
        ]  
      },  
      {  
        "name": "practical_implementation_guide.md",  
        "type": "Код & Реализация",  
        "pages": 30,  
        "topics": [  
          "Полный PyTorch код MoE",  
          "Expert и Gating слои",  
          "Training loop с мониторингом",  
          "Инференс pipeline",  
          "Оптимизации для продакшена"  
        ]  
      },  
      {  
        "name": "final_metrics_and_recommendations.md",  
        "type": "Анализ & Рекомендации",  
        "pages": 40,  
        "topics": [  
          "Сравнительные таблицы",  
          "Анализ экономии ресурсов",  
          "Сценарии использования",  
          "Гиперпараметры и чувствительность",  
          "Чек-листы деплоя"  
        ]  
      },  
      {  
        "name": "quick_reference_cheatsheet.md",  
        "type": "Быстрая справка",  
        "pages": 2,  
        "topics": [  
          "Архитектура в одном графике",  
          "Параметры компонентов",  
          "Ключевые метрики",  
          "Отладка проблем",  
          "Quick start код"  
        ]  
      },  
      {  
        "name": "complete_integrated_pdf.md",  
        "type": "Полный отчёт",  
        "pages": 120,  
        "topics": [  
          "Части 1-4 объединены",  
          "Приложения с диаграммами",  
          "Детальная техническая документация",  
          "Аналитические отчеты и рекомендации"  
        ]  
      }  
    ]  
  }  
}
```

Автоформатировать

```
    Структурирован для читирования
    ]
  },
{
  "name": "index_and_guide.md",
  "type": "Индекс & Руководство",
  "pages": "20+",
  "topics": [
    "Структура пакета",
    "Содержание всех документов",
    "Чек-листы и сценарии",
    "FAQ и советы",
    "Ссылки на ресурсы"
  ],
}
],
},
"key_metrics": {
  "dense_baseline": "20B параметров",
  "moe_system": "11.8B параметров",
  "improvements": {
    "parameter_efficiency": "-41%",
    "flops_reduction": "-88%",
    "memory_savings_weights": "-41%",
    "memory_savings_activations": "-75%",
    "training_speedup": "4x",
    "inference_speedup": "3.3x",
    "expansion_cost_reduction": "-87%",
    "accuracy_drop": "-0.7% (acceptable)"
  }
},
"recommended_usage": {
  "best_for_moe": [
    "5-100+ специализированных задач",
    "Требуется быстрая расширяемость",
    "Ограниченнная GPU память (1-2 A100)",
    "Допустима потеря ±1% точности",
    "Критично время обучения"
  ],
  "use_dense_if": [
    "Только 1-3 задачи",
    "Требуется максимальная точность",
    "Budget GPU неограничен",
    "Приоритет - простота отладки"
  ]
},
"phases": {
  "phase_1_prototyping": {
    "duration": "2 недели",
    "tasks": [
      "Базовая архитектура",
      "Валидация данных",
      "Baseline метрики"
    ]
  },
  "phase_2_optimization": {
    "duration": "2 недели",
    "tasks": [
      "Гиперпараметры",
      "Aux loss оптимизация",
      "A/B тестирование"
    ]
  },
  "phase_3_scaling": {
    "duration": "4 недели",
    "tasks": [
      "Добавить ветки",
      "Continuous learning",
      "Интеграция с CI/CD"
    ]
  }
}
```

Автоформатировать

```
        },
        "phase_4_production": {
            "duration": "ongoing",
            "tasks": [
                "Деплой",
                "Метрики",
                "Улучшение"
            ]
        }
    },
    "quick_start": {
        "installation": "pip install torch transformers peft tensorboard wandb",
        "model_init": "model = InterconnectedAISystem(num_experts=8, k=2, num_branches=3)",
        "training_time_per_epoch": "13.9 часов (vs 55.6 Dense)",
        "gpu_requirements": "1x A100 80GB"
    }
}
```