Computer Use Safety Documentation

This document outlines the safety features and considerations for the voice assistant's computer use capabilities.

Safety Levels

Off Level

- **Description**: Computer use completely disabled
- **Capabilities**: None
- **Use Case**: Maximum security, conversation only
- **Risk Level**: None

Safer Level (Recommended)

- **Description**: Limited safe computer operations
- **Capabilities**:
- File reading in user directories
- System information queries
- Process listing
- Safe file operations in Documents, Downloads, Desktop
- **Restrictions**:
- No system file access
- No administrative operations
- User confirmation required
- All actions logged
- **Use Case**: General productivity with safety
- **Risk Level**: Low

God Level (Dangerous)

- **Description**: Full system access
- **Capabilities**: Everything the user can do
- **Restrictions**: Only blocked commands list
- **Use Case**: Advanced automation (experts only)
- **Risk Level**: High - can damage system

Blocked Commands

The following commands are always blocked regardless of safety level:

```
rm -rf /                    # Delete entire filesystem
dd if=/dev/zero             # Disk destruction
:(){ :|:& };:               # Fork bomb
chmod -R 777 /              # Dangerous permissions
mkfs                        # Format filesystem
fdisk                       # Disk partitioning
```

Safe Directories

In “safer” mode, file operations are restricted to:
- `~/Documents`
- `~/Downloads`
- `~/Desktop`
- `~/Pictures`
- `~/Music`
- `~/Videos`
- `/tmp` (read only)

Sensitive Paths

These paths are always protected:
- `/etc/passwd` - User accounts
- `/etc/shadow` - Password hashes
- `/etc/sudoers` - Sudo configuration
- `/boot` - Boot files
- `/sys` - System files
- `/proc` - Process information

User Confirmation

In safer and god modes, the following operations require user confirmation:
- File modifications outside safe directories
- System command execution
- Process termination
- Network operations
- Application launching

Action Logging

All computer use actions are logged with:
- Timestamp
- Action type
- Command/operation details
- Safety level at time of action
- Success/failure status

Logs are stored locally and can be reviewed for security auditing.

Best Practices

For Regular Users

1. **Use “safer” mode** for daily operations
2. **Review confirmation prompts** carefully
3. **Check action logs** periodically
4. **Never enable “god” mode** unless absolutely necessary

For Advanced Users

1. **Understand the risks** of god mode
2. **Have backups** before enabling full access
3. **Monitor system activity** during automation
4. **Use time-limited sessions** for god mode

For Developers

1. **Test in safer mode first**
2. **Implement additional safety checks**
3. **Validate all user inputs**
4. **Use principle of least privilege**

Security Considerations

Potential Risks

- **Accidental file deletion** in god mode
- **System configuration changes** through voice commands
- **Unintended command execution** due to speech recognition errors
- **Privilege escalation** if running as administrator

Mitigation Strategies

- **Default to safer mode** on installation
- **Require explicit confirmation** for dangerous operations
- **Implement command validation** and sanitization
- **Maintain comprehensive logging** for audit trails
- **Regular safety rule updates** for new threats

Emergency Procedures

If Something Goes Wrong

1. **Stop the assistant immediately**: Ctrl+C or close application
2. **Check action logs**: Review what commands were executed
3. **Assess damage**: Check file system and system state
4. **Restore from backup**: If files were modified or deleted
5. **Report issues**: Document problems for improvement

Recovery Commands

```
# Stop all assistant processes
pkill -f voice_assistant

# Check recent file modifications
find ~ -mtime -1 -type f

# Review system logs
journalctl --since "1 hour ago"

# Check for unusual processes
ps aux | grep -v "^\[.*\]$"
```

Configuration Examples

Maximum Security

```
{
  "computer_use": {
    "safety_level": "off",
    "require_confirmation": true,
    "log_actions": true
  }
}
```

Balanced Security

```
{
  "computer_use": {
    "safety_level": "safer",
    "require_confirmation": true,
    "log_actions": true,
    "allowed_commands": ["file_operations", "system_info"]
  }
}
```

Development/Testing (Risky)

```
{
  "computer_use": {
    "safety_level": "god",
    "require_confirmation": true,
    "log_actions": true,
    "log_all_actions": true
  }
}
```

Compliance and Legal

Data Protection

- All operations are performed locally
- No data is transmitted to external services

- User maintains full control over all data

Liability

- Users are responsible for actions performed by the assistant
- Safety features are provided as-is without warranty
- Always test in safe environments first

Audit Requirements

- Action logs provide complete audit trail
- Timestamps and details for all operations
- Exportable logs for compliance reporting

Updates and Maintenance

Safety Rule Updates

- Regularly update blocked command lists
- Add new sensitive path protections
- Review and improve validation logic

Security Patches

- Monitor for new security vulnerabilities
- Update dependencies regularly
- Test safety features after updates

User Education

- Provide clear documentation
- Warn about risks appropriately
- Offer training for advanced features

Conclusion

The computer use safety system is designed to provide useful automation while minimizing risks. However, no system is perfect, and users must understand the implications of enabling computer control through voice commands.

Remember: With great power comes great responsibility. Use computer use features wisely and always maintain proper backups.

Emergency Contact: If you discover a security vulnerability, please report it immediately through the project’s security channels.